

# Safe but sorry: Theory, assessment, and treatment of pain-related fear in chronic low back pain patients

Citation for published version (APA):

Leeuw, M. (2008). *Safe but sorry: Theory, assessment, and treatment of pain-related fear in chronic low back pain patients*. [Doctoral Thesis, Maastricht University]. Datawyse / Universitaire Pers Maastricht. <https://doi.org/10.26481/dis.20080307ml>

## Document status and date:

Published: 01/01/2008

## DOI:

[10.26481/dis.20080307ml](https://doi.org/10.26481/dis.20080307ml)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

# **SAFE BUT SORRY**

**Theory, assesment, and treatment  
of pain-related fear in chronic low back pain patients**

SAFE but SORRY: Theory, assessment, and treatment of pain-related fear in chronic  
low back pain patients  
Maaïke Leeuw

© Copyright Maaïke Leeuw, Maastricht 2008

Universitaire Pers Maastricht  
ISBN 978-90-5278-696-4

Omslagontwerp & layout: Maaïke Leeuw | Datawyse  
Druk: Datawyse

# **SAFE BUT SORRY**

**Theory, assesment, and treatment  
of pain-related fear in chronic low back pain patients**

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit Maastricht,  
op gezag van de Rector Magnificus, Prof. Mr. G.P.M.F. Mols,  
volgens het besluit van het College van Decanen,  
in het openbaar te verdedigen op vrijdag 7 maart 2008 om 16:00 uur

door

Maaïke Leeuw



**PROMOTOR**

Prof. Dr. J.W.S. Vlaeyen

**COPROMOTOR**

Dr. M.E.J.B. Goossens

**BEOORDELINGSCOMMISSIE**

Prof. Dr. A.R. Arntz (voorzitter)

Prof. Dr. F.J. Keefe

Dr. J. Roelofs

Prof. Dr. O. Van den Bergh

Dr. J.R. Verbunt

## TABLE OF CONTENTS

<b>Part I: Introduction</b>	7
<b>Chapter 1</b> General introduction	9
<b>Part II: Outlining and testing the contribution of pain-related fear in chronic low back pain</b>	25
<b>Chapter 2</b> The fear avoidance model of musculoskeletal pain: Current state of scientific evidence	27
<b>Chapter 3</b> Pain-related fear in low back pain: A prospective study in the general population	47
<b>Part III: Assessment of pain-related fear</b>	63
<b>Chapter 4</b> Fear of movement/injury in the general population: Factor structure and psychometric properties of an adapted version of the Tampa Scale for Kinesiophobia	65
<b>Chapter 5</b> Measuring perceived harmfulness of physical activities in patients with chronic low back pain: The Photograph Series of Daily Activities – Short electronic Version	79
<b>Chapter 6</b> Measuring fear of movement/(re)injury in chronic low back pain using implicit measures	95
<b>Part IV: Clinical implications of pain-related fear for chronic low back pain management</b>	109
<b>Chapter 7</b> Exposure in vivo versus operant graded activity in chronic low back pain patients: Results of a randomized controlled trial	111
<b>Chapter 8</b> Assessing protocol adherence and contamination in treatment outcome studies: A methodological illustration from behavioural medicine	141

<b>Part V: Summary &amp; Discussion</b>	159
<b>Chapter 9</b>	
Summary & General Discussion	161
<b>Nederlandse samenvatting</b> (Dutch Summary)	189
<b>References</b>	199
<b>Dankwoord</b>	223
<b>Curriculum Vitae</b>	231
<b>Publications</b>	235

# Part I

## Introduction





# **Chapter 1**

## **General introduction**



### **The problem of low back pain**

Nearly everyone will experience low back pain (LBP) complaints at least once during his/her life. From large population based studies, the point prevalence of LBP is estimated to be 15-30%, the one-month prevalence 19-43%, and the lifetime prevalence between 60-70% (Nachemson et al., 2000), or even up to 80% (Walker, 2000). In the Netherlands, 44% of the people indicated to have experienced LBP in the past year, and 27% at the current moment, whereas 21% reported to have chronic complaints (lasting longer than three months) at that time (Picavet and Schouten, 2003). Most people experience their first episode of LBP between the age of 20 and 55 years.

LBP is defined as pain in the area of the lower back, which in most people also radiates to the buttocks or the legs. According to the International Association for the study of Pain, pain is defined as 'An unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage', indicating that pain is not necessarily associated with biomedical pathology. Indeed, the majority of patients (90%) present with ordinary backache for which no specific biomedical cause can be found, which is denoted as *non-specific* LBP. The remaining 10% of the patients are diagnosed with *specific* LBP, since either nerve root pain or a serious spinal disease underlies their pain problem (Waddell, 2004b). Although LBP is classified into an acute (less than six weeks), subacute (six to twelve weeks), or chronic (more than twelve weeks) condition (Faas et al., 1996), epidemiological studies have shown that these complaints often follow a recurrent, intermittent, and fluctuating pattern, with previous experiences of LBP as the strongest predictor for the occurrence of a new episode (von Korff et al., 1988; Linton and Ryberg, 2000; Nachemson et al., 2000; Picavet and Schouten, 2003).

### **The consequences of low back pain**

Although it is not hard to imagine that pain affects people in their daily life activities, its consequences vary widely between patients. Pain severity and functional disability are not utterly related to each other, with some patients still being able to function normally despite severe pain, and other patients with only mild pain experiencing limitations in daily living, and/or seeking medical care for their complaints. Despite these individual differences, in general, LBP is associated with substantial perceived disability, work absenteeism, and health care utilization (Picavet and van den Bos, 1997; Linton and Ryberg, 2000; Picavet and Schouten, 2003).

LBP not only influences the affected patients and their social environment, but it also has a major impact on society. Whereas the prevalence and presentation of LBP have not changed over time, its socioeconomic consequences have increased considerably in the past few decades, especially due to LBP associated work absenteeism and disability benefits (Nachemson et al., 2000; Goossens, 2002; Waddell, 2004c). Musculoskeletal diseases (of which 50% LBP) constitute the most expensive disease category in the Netherlands with regard to disability and absence from work (van Tulder et al., 1995). Additionally, together with mental problems, musculoskeletal complaints account for the highest health care costs (Meerding et al., 1998). In fact, LBP places a large economic burden on society, since its total costs are estimated around 1-2% of the Gross National Product in three Western countries (Norlund and Waddell, 2000). The largest part (90%) of these comprise indirect costs,

such as sick leave and disability benefits, whereas the direct costs, such as health care utilization, only account for the remaining 10% (Norlund and Waddell, 2000; Goossens, 2002). The majority of these costs are accounted for by the small group of patients that suffer from chronic complaints (Picavet et al., 1999; Linton and Ryberg, 2000; Maetzel and Li, 2002).

Since the individual and societal consequences of LBP are the most severe in chronic low back pain (CLBP) patients, it may be worthwhile to intervene on, and aim to reduce the disabling consequences of LBP, in these chronic patients.

### **A bio-psycho-social view of low back pain**

For the understanding and management of pain conditions, it is important to distinguish acute from chronic pain. *Acute* pain is typically associated with nociception or tissue damage, and the pain and disability are mainly in accordance with physical findings. Acute pain can therefore be successfully managed with biomedical care. In contrast, in *chronic* pain, the biomedical model that focuses upon structural and biomechanical abnormalities as a prerequisite for pain and associated consequences, no longer adequately explains the experience of pain and disability (Mannion et al., 2001; Heuts et al., 2004; Waddell, 2004a). In this stage of pain, resting and avoidance of painful movements aggravate the complaints instead of being effective, and biomedical treatments seem futile. A bio-psycho-social view has been introduced to explain these chronic pain complaints. According to this model, pain and disability are not only determined by biomedical pathology, but also by the patient's beliefs, coping strategies, and distress (psychological factors), and his/her behavioural response to the complaints, as well as social and cultural interactions (social factors) (Waddell, 2004a).

### **Pain-related fear and low back pain**

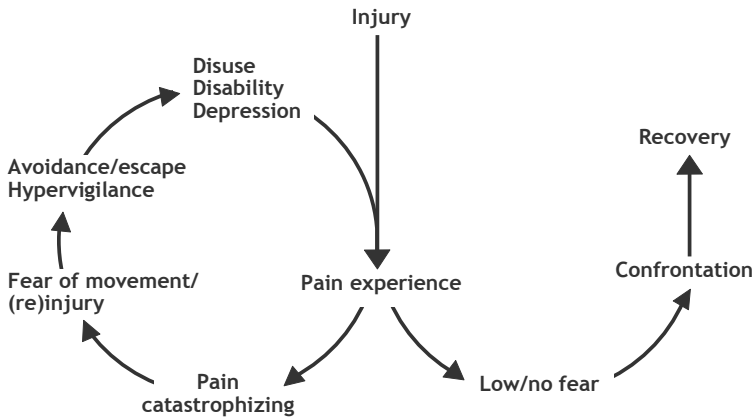
One psychological factor that has been the subject of many research studies with regard to the development and maintenance of CLBP, is fear. Pain patients may fear the causes and consequences they assume to be associated with their pain. Depending on the source to which the fear is directed, patients may for example fear the unpleasantness of the pain experience, the performance of movements, or the (re)occurrence of injury. Pain-related fear is therefore a collective term, under which several specific fears can be shared, such as *fear of movement/(re)injury* that specifically denotes the fear that certain movements may cause (re)injury to the back, and the strongly associated construct of *perceived harmfulness of activities*, signifying the degree to which performing various activities is believed to be harmful to the back.

The fear avoidance model of chronic musculoskeletal pain (Vlaeyen et al., 1995a; Vlaeyen and Linton, 2000) assigns a prominent role to pain-related fear in the development and maintenance of CLBP. According to this model, the manner in which an acute pain experience is appraised and responded to, will determine whether someone will either recover or develop enduring complaints. The adaptive response consists of the appraisal of pain as unpleasant but non-threatening, and the preservation of adequate and sufficient activity encounters, due to which healing and recovery are promoted. Fortunately, most people respond in this adaptive manner to an acute LBP episode. On the other hand, the minority of people regard their pain as a threatening experience, and may believe that their back is seriously damaged and

requires careful protection (pain catastrophizing). Due to these catastrophic cognitions about the causes and consequences of their pain, these people may become afraid of pain, movement, and aggravation of their injury. This pain-related fear may subsequently give rise to safety seeking behaviours, such as activity avoidance/escape and excessive attention to pain and pain-related information, which can be adaptive in the acute pain stage, but paradoxically worsen the problem in case of long-lasting pain. Specifically, continuous avoidance of activities may result in disability, social withdrawal, depressed mood, and a deterioration of physical fitness, which in turn may perpetuate the pain problem (Figure 1). Many studies have provided support for several proposed interrelationships of the fear avoidance model, although the latest review dates from 2000 (Vlaeyen and Linton, 2000).

**Figure 1**

The fear avoidance model of musculoskeletal pain (adapted from: Vlaeyen JWS & Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: A state of the art. *Pain* 2000;85(3):317-332)



Pain-related fear is shown to exist in CLBP patients (e.g. Crombez et al., 1999b) and acute LBP patients (e.g. Sieben et al., 2005a; Swinkels-Meewisse et al., 2006c). But also in pain-free people, dysfunctional beliefs about pain and its consequences have been documented, such as ‘when one has back pain, a wrong movement can lead to serious problems’ (e.g. Picavet et al., 2002; Goubert et al., 2004a). Moreover, pain-related fear is found to be associated with the initiation and continuation of LBP (e.g. Picavet et al., 2002; Swinkels-Meewisse et al., 2006c), and it might therefore be an important target for intervention in some CLBP patients.

### Assessment of pain-related fear

With pain-related fear being such an important factor in the development and maintenance of (C)LBP, it is useful to have reliable and valid measures to be able to determine a patient’s degree of fear. At present, several major self-report measures are available for this purpose (see also McNeil and Vowles, 2004), which are the Fear Avoidance Beliefs Questionnaire (Waddell et al., 1993), the Fear of Pain Questionnaire III (McNeil and Rainwater, 1998), the Pain Anxiety Symptoms Scale

(McCracken et al., 1992), and the Tampa Scale for Kinesiophobia (TSK: Kori et al., 1990). Despite the availability of these adequate self-report measures for the assessment of pain-related fear, it is worth the effort to develop new ones that have additional value to these.

#### *Assessment of fear of movement/(re)injury in pain-free people*

Dysfunctional LBP beliefs in pain-free people have been associated with subsequent inceptions of LBP (Linton et al., 2000; Picavet et al., 2002; Linton, 2005). These pain beliefs may result from previous LBP encounters, from general views of LBP that prevail in the population, or from personality traits. Psychometrically sound instruments are therefore welcomed to assess these beliefs in currently pain-free people. For this purpose, a modified version of the TSK (the TSK-G: Vlaeyen and Crombez, 1998) has been utilized, in which the items of the TSK were adapted in such way that pain-free people were also able to complete it (Peters et al., 2002; Picavet et al., 2002). Despite the TSK-G holding promise in the assessment of fear of movement(re)injury in pain-free people, its psychometric properties have never been studied.

#### *Assessment of perceived harmfulness of activities*

It can be argued that the specific worries of certain CLBP patients might concern the perceived harmfulness of activities, that is their belief that certain activities can be harmful or damaging to the back. One instrument that specifically assesses these beliefs is the Photograph Series of Daily Activities (PHODA: Kugler et al., 1999). It consists of 100 photographs of daily activities that patients are requested to arrange in increasing order along a harmfulness thermometer according to the extent that they consider these activities to be harmful to their back. The PHODA has been applied as a clinical tool in an exposure in vivo treatment as well as a measurement instrument in CLBP patients (e.g. Vlaeyen et al., 2002b; Boersma et al., 2004; de Jong et al., 2005b), but its psychometric properties are yet to be investigated. Moreover, several aspects of the PHODA are open to improvement, since its application is fairly time-consuming and not completely standardized, and because its data need to be manually stored.

#### *Implicit assessment of fear of movement/(re)injury*

Most measures assess pain-related fear by means of patients' self-report. However, there may be several disadvantages attached to these kind of measures (de Houwer, 2002; de Houwer, 2003b; Fazio and Olson, 2003). First, they require the patients' ability to consciously reflect upon the construct being measured. For fear of movement/(re)injury, patients need to be aware of their fear and able to identify and designate the content of their worries. In CLBP patients this might not always be the case (Vlaeyen et al., 2002c; Vlaeyen et al., 2004a). Second, since it is obvious to the patients what the measures aims to assess, their responses might be influenced by biases, such as self-presentational strategies. For example, when assessed before and after treatment, patients may (consciously or unconsciously) exaggerate the decrease in their scores since they expect this to be in line with the therapists' or researchers' prospects. It is therefore worthwhile to search for additional assessment methods that overcome these limitations of self-report measures.

In the past few years, there is a major rise in the development and application of implicit measures, which aim to measure evaluations or associations that are automatically activated without the need for awareness or control of the participant (e.g. Greenwald and Nosek, 2001; Nosek and Banaji, 2001; de Houwer, 2003a). These often consist of reaction time paradigms during which a response is required to words that are related to the association of interest. Implicit associations are assumed to be reflected in facilitated task performance on trials that are compatible to the underlying association, and in hampered performance on trials that are incompatible with the presumed association (e.g. de Houwer, 2003b). For example, it should be easier to press a positive key in response to flowers than to bugs, because of a universal positive attitude towards the former and not the latter construct. Two previous studies already aimed at assessing pain-related fear in CLBP patients, demonstrating ambiguous results thus far (de Jong and Peters, 2002; Goubert et al., 2003). Since neither of these studies assessed the association between movements and the evaluation threatening, which is the association we believe underlies fear of movement/(re)injury, it might be worthwhile to extend on these previous efforts in assessing pain-related fear in an implicit manner.

### **Two theoretical models of low back pain: when learning can be dysfunctional**

Before examining how the management of CLBP, and specifically the two interventions that are relevant within the scope of this dissertation, may be designed, first the two theoretical models that underlie these interventions are outlined. Both models stress the importance of learning in people. Most importantly, people need to learn which stimuli or events are safe, and which are not, to optimize their chance of survival. Learning takes place by encountering consequences upon acting (learning through operant conditioning), and by discovering associations between simultaneous or successive events (learning through classical conditioning). According to these models, chronic pain can be developed and maintained by certain learning processes that generally are adaptive, but may have become dysfunctional in the case of chronic pain.

#### *Operant conditioning and pain*

According to the operant learning principles, people learn to operate upon their environment according to the consequences of their behaviour. Behaviour followed by a positive consequence, which is either the presence of something positive (*positive reinforcement*) or the prevention or removal of something negative (*negative reinforcement*), will increase in frequency or intensity and will be more likely to recur. On the other hand, when behaviour has negative consequences, either because something negative occurs (*positive punishment*) or something positive is ended or taken away (*negative punishment*), that specific behaviour will decrease in frequency or intensity and will be less likely to recur (Sanders, 2002).

These operant learning principles were first applied in the area of chronic pain in 1976 by Fordyce. In this view, the focus is on observable pain behaviours, such as verbal (complaining) and nonverbal pain expressions (groaning, body postures, facial expressions), or medication or health care use. As any other behaviour, these pain behaviours are learned and maintained by their consequences. There are four operant mechanisms that may contribute to maladaptive learning in CLBP patients. *First*,



when pain behaviour is positively reinforced by eliciting favourable consequences, for example spousal attention and empathy or compensation benefits in case of work incapacity, this behaviour may increase in duration or frequency, and is thereby enhanced and maintained. *Second*, pain behaviour can be negatively reinforced when it involves avoidance of negative consequences, such as stressful interpersonal situations, or undesirable home and work responsibilities. But probably more important, patients may learn to avoid activities that they expect to be painful or damaging, the so-called 'avoidance learning'. The decrease or non-occurrence of pain or fear consequently to the avoidance behaviour negatively reinforces this behaviour. The *third* operant mechanism of influence for chronic pain is the deficiency of reinforcing healthy behaviour (e.g. staying active despite pain). This is especially the case when the patient is convinced that it is not advisable or possible to perform certain activities, which may also be corroborated by significant others and therapists. *Fourth*, in addition to lack in reinforcement of healthy behaviour, actual active behaviour may be followed by positive punishment, that is intensified pain, and therefore discouraged to recur. Due to the contingencies upon pain behaviours, these may develop into operant responses that eventually act independent of the level of biomedical pathology or nociception (Fordyce, 1976; Fordyce, 2001b; Fordyce, 2001a; Sanders, 2002; Goubert et al., 2004b).

Thus far, only indirect support is found for operant learning mechanisms in the development and maintenance of pain. Preliminary support was provided by studies showing that pain behaviour can be influenced by operant contingencies (Fordyce et al., 1968). Already in 1977, Cairns and Pasino showed that, in contrast to a control condition without any reinforcement, activity levels of CLBP patients were significantly enhanced by verbal reinforcement of activity increases. Also pain intensity reports may be operantly controlled by verbal reinforcement (e.g. White and Sanders, 1986). Experimental studies with healthy participants undergoing painful stimuli showed that pain intensity reports either increased or decreased depending on whether the verbal reinforcement was delivered upon heightened or diminished pain report respectively, as compared to the control condition without any reinforcement, and that this was even found when the painful stimulus actually diminished in intensity (Linton and Gotestam, 1985; Chambers et al., 2002; Flor et al., 2002; Jolliffe and Nicholas, 2004). The same effect was shown with regard to facial pain expressions (Chambers et al., 2002), and physiological reactivity (Lousberg et al., 1992; Flor et al., 2002). Of specific interest was the finding of Flor et al. (2002) that CLBP patients showed slower extinction of operantly acquired behaviour, suggesting that these patients may be more susceptible to operant learning than healthy people. Further support for the operant model of chronic pain is provided by studies that focussed upon spousal influences on pain behaviour. These studies have found that the presence of a solicitous spouse substantially increased verbal pain reports (Block et al., 1980), non-verbal pain behaviour (Romano et al., 2000), and perceived limitations in functioning (Schwartz et al., 2005), and that it decreased walking duration (Lousberg et al., 1992), and pain tolerance levels (Flor et al., 1995). Furthermore, patients demonstrated higher activity levels when their pain was ignored or negatively responded to by their spouses (Flor et al., 1987).

Since the emphasis of these previous studies is on pain reports and behaviour rather than on the pain experience itself, and the participants may have been aware of

the contingencies delivered, a recent study employed an operant learning paradigm that could intrinsically reinforce changes in pain stimulation itself without the awareness of the participant. With this method, it was shown that short-term sensitization as well as habituation to thermo-nociceptive stimuli could be shaped in healthy subjects. This may explain the operant learning of hypersensitivity, in that the negative reinforcement of avoidance behaviour due to pain reduction may increase a chronic pain patient's sensitiveness to painful stimuli (Holzl et al., 2005).

### *Classical conditioning and pain*

In addition to learning a stimulus-consequence association as is the case in operant conditioning, people may also learn associations between successive events, which is the case in classical conditioning. The classical example illustrating the principle of this form of conditioning is the Pavlovian dog that already started to salivate when hearing a bell ringing even in the absence of any food, when before that the food was repeatedly presented after the ringing of a bell. Thus, classical conditioning proposes that, whereas at first an unconditioned response (e.g. salivating) follows an unconditioned stimulus (e.g. the food), after repeated encounters of a neutral stimulus (e.g. the ringing bell) preceding this unconditioned stimulus, the neutral stimulus by itself can eventually elicit a conditioned response that resembles the unconditioned response (salivating in response to the ringing bell), and thus becomes a conditioned stimulus. Applied to CLBP, when patients repeatedly experience pain increases as a result of certain activities, eventually the mere prospect of these activities revives the association and expectations of pain increase, and will result in avoidance and escape behaviours. This avoidance behaviour is inevitably encouraged and sustained, since these behaviours prevent patients from discovering that the presumed detrimental or painful consequences are actually non-existent (Vlaeyen and Linton, 2000).

Several recent insights in the process of classical conditioning shed a new light on learning, of which only those relevant to this dissertation will be discussed with respect to CLBP (Goubert et al., 2004b). All of these built upon the proposition that learning is more dynamic than lead to believe by the historical conditioning perspective, and that cognitive aspects are of influence as well. *First*, CLBP patients do not necessarily have to learn to associate certain activities with pain due to their own experiences, but the expectation of impending pain can also originate from beliefs of activity-pain associations in the environment, for example expressed by the therapist or significant others. *Second*, the aversiveness of the unconditioned stimulus (e.g. pain) can vary across patients according to personal characteristics. For example, patients who catastrophize about the pain experience will interpret their pain as more threatening and unpleasant than those low in catastrophizing, who experience it as a minor temporal discomfort. Furthermore, the aversiveness of the unconditioned stimulus 'pain' can become inflated even in the absence of the original conditioned stimulus, accordingly resulting in more intense anticipatory fear responses. For example, even though a CLBP patient has only experienced mild pain during lifting, a sudden pain attack in the absence of lifting can intensify subsequent fear of lifting. Likewise, observing another person becoming severely disabled by LBP may increase pain-related fear. The same mechanisms apply for socially or verbally transmitted information, such as when patients are warned by their healthcare providers not to engage in lifting because this can aggravate their already damaged and painful back.

*Third*, there is a clear asymmetry between the ease in which acquisition occurs as compared to the difficulty by which extinction takes place. It is difficult, or maybe even impossible, to unlearn an acquired association between the conditioned and unconditioned stimulus. Extinction should rather be viewed as additional learning that in a particular context the association does not exist. For example, when a fearful CLBP patient successfully performs a lifting task, this patient may rather perceive this as an exception to the presumed association between lifting and damage, than that the association is reduced or abolished (Crombez et al., 2002b; Goubert et al., 2002; Goubert et al., 2005a). It may therefore be very difficult to generalize the extinction of fear to other dissimilar contexts, making relapse likely to occur (Bouton, 2002; Mineka and Zinbarg, 2006).

With regard to CLBP, classical conditioning may apply to muscle tension as well as the pain experience, although hardly any studies examined the latter (Goubert et al., 2004b). While muscle tension surrounding the injured or painful area is proposed to be an instinctive reaction, possibly because it may promote recovery due to enhanced immobility, long-lasting muscle tension may actually aggravate the pain experience. Some studies indeed showed anomalous muscle tension of the back muscles in CLBP patients, in that these patients specifically respond with hyperreactivity of the paraspinal muscles to stressful situations (Flor et al., 1985; Flor et al., 1991; Peters and Schmidt, 1991; Flor et al., 1992; Burns, 1997). These studies do not demonstrate, however, that muscle tension is a conditioned response. With respect to this, Schneider et al. (2004) showed that, as compared to healthy controls, CLBP patients responded with enhanced muscular reactivity of their left arm to an aversive slide, when before that the aversive slide repeatedly preceded an intracutaneous electric stimulus at that arm. Furthermore, CLBP patients were less susceptible to unlearn this conditioned response. Comparable results were already found by Flor and Birbaumer (1994) with students suffering from sub-chronic pain.

#### *A combination of both models: the two-factor model of avoidance learning*

According to the two-factor theory of Mowrer (1947), both classical conditioning and operant conditioning are involved in avoidance learning. First, fear is acquired by classical conditioning, whereas the conditioned fear is thereafter maintained by operant conditioning. Specifically, the avoidance behaviour is presumed to be negatively reinforced, and thereby maintained, by the reduction or elimination of fear. Moreover, the interaction between Pavlovian and operant processes is dynamic and continuously changing. However, there are some challenges to this model. For example, persistent avoidance responding in the absence of aversive events should eventually result in extinction of fear, and subsequent diminishment of the operant avoidance response, but this does not seem to be the case (Domjan, 2005).

### **Clinical management of chronic low back pain**

Health care providers have to deal with certain challenges when confronted with CLBP patients seeking medical care. Besides the intricacy of diagnosing non-specific CLBP (Nachemson and Vingard, 2000), clinicians also face with shortage in effective treatments that can provide the pain relief these patients are yearning for. Recently, rehabilitation treatments have therefore adopted the bio-psycho-social view of chronic pain, with restoration of patients' functional abilities, and not pain reduction, as

treatment goal. These treatments incorporate various cognitive behavioural techniques, aiming to increase patients' activity levels despite pain, to decrease illness behaviours and increase healthy behaviour, to change maladaptive cognitions about pain, and to enhance self-control and acceptance of pain. Usually, rehabilitation treatments are carried out in a multidisciplinary manner, with rehabilitation physicians, physical therapists, occupational therapists, and psychologists concurrently contributing to the treatment of the patient. Meta-analyses generally found that (cognitive) behavioural treatments are effective as compared to a waiting list or no treatment control, but that these effects are only moderate to small, and that the additional value of these approaches contrasted to other active interventions may be even smaller. Furthermore, there is no specific behavioural treatment approach that excels the others (van Tulder et al., 1997; Morley et al., 1999; Guzman et al., 2001; van Tulder et al., 2001; Hofmann et al., 2005; Ostelo et al., 2005).

The two multidisciplinary (cognitive) behavioural treatments that are subject of this dissertation are based on the previously described theoretical models of learning with respect to chronic pain, and will now be described in more detail.

#### *Operant conditioning: graded activity*

Treatments that are based on the operant model of pain generally do not strive for pain relief, but rather emphasize the identification of reinforcement contingencies that contribute to the persistence of enduring pain, and subsequently aim to reverse these. Additionally, healthy behaviours (e.g. being active despite pain) are encouraged. The most prominent and applied operant treatment is graded activity (GA) that combines such an intervention strategy with a time-contingent increase of activities (Fordyce, 2001b; Fordyce, 2001a; Sanders, 2002; Geilen et al., 2005). First, the baseline level of activities is determined by letting the patient perform various activities within his/her limits of tolerance. The patient is thus requested to function in a pain contingent manner, which is generally reflected in a fluctuating activity pattern with increased activity on 'good days' and decreased activity on 'bad days'. Then, a treatment plan is generated, in which activities should be initially performed somewhat below the baseline activity level, after which the activity quota are gradually increased and appointed in time towards preset treatment goals wished for by the patient. When, for example, a patient endeavours to be able to walk for one hour, whereas at baseline on average he is only able to walk twelve minutes, according to the time-contingent treatment plan his walking duration should start with ten minutes of walking, which should then be increased with five minutes each week during a period of ten weeks. Due to these time-contingent activity agreements, patients learn to be active in a time-contingent rather than a pain-contingent manner, due to which the negative reinforcement of pain relief associated with resting and avoidance no longer takes place. Furthermore, the accomplishment of these appointed treatment quota (healthy behaviour) is positively reinforced by providing idiosyncratic rewards identified for that patient. At first, this positive reinforcement of activity increase is consequently administered, but later on this only takes place intermittently, to strengthen the acquisition of the newly learned behaviour. In the mean time, the therapists are unresponsive to patients' pain behaviour such as grimacing and verbal complaints about pain, so these behaviours are no longer reinforced. Finally, when relevant, the influence of the spouse on the patient's pain behaviour is integrated in the treatment.

In sum, pain behaviour, such as resting and complaining is no longer reinforced, whereas healthy behaviour is promoted, thereby increasing patients' activity levels and decreasing functional disability despite pain. Various studies corroborated the effectiveness of GA in improving functional abilities in CLBP patients (Lindstrom et al., 1992; Kole-Snijders et al., 1999; van Tulder et al., 2001).

*Classical conditioning: exposure in vivo*

Treatments that are based on the classical conditioning model of pain do not strive for pain relief either, but aim at extinguishing fear and the presumed associations of patients between certain activities and consequent damage. Because of many remarkable similarities between pain-related fear and fear in general, together with successful applications of exposure in vivo (EXP) to feared stimuli in phobic patients, an EXP treatment was developed for application in fearful CLBP patients. First, the patients' specific worries about the causes and consequences of pain and activities are elaborated on by a psychological intake. Then, an idiosyncratic fear hierarchy is established by use of the PHODA, consisting of activities that are increasingly associated with damage to the back by that patient. Then, gradual EXP exercises are developed based on this fear hierarchy, starting with exposing patients to activities that induce only mild fear and subsequently moving forward to activities that the patients consider as highly threatening. Generally, because of a more cognitive emphasis on learning, behavioural experiments are added to these EXP practices, during which patients can test the validity of their beliefs about the harmful consequences of activities, and can construct and test new, more adaptive beliefs. Due to these EXP exercises, patients can experience that the expected detrimental consequences are in fact catastrophic overestimations. Consequently, extinction of the association between activities and damage takes place, catastrophic cognitions are rectified, fear is diminished, and functional abilities are promoted (Vlaeyen et al., 2002c; Vlaeyen et al., 2004a). For example, a patient afraid of lifting because of the conviction that this may damage a nerve, may be encouraged to perform this specific activity to learn that he/she is capable of lifting without inducing nerve damage. From recent insights in classical conditioning learning, EXP is practiced within various contexts, with varying stimuli, and is spaced over time, to enhance generalization of the newly learned associations that activities can be performed without inducing or aggravating damage to the back (Crombez et al., 2002b; Goubert et al., 2002; Goubert et al., 2005a). The effectiveness of EXP combined with behavioural experiments as compared to GA or a no-treatment baseline period in highly fearful CLBP patients has been thus far supported by several studies using replicated single case experimental designs (Vlaeyen et al., 2001; Linton et al., 2002a; Vlaeyen et al., 2002a; Vlaeyen et al., 2002b; Boersma et al., 2004; de Jong et al., 2005b).

*Which treatment works for whom, and why?*

Although cognitive behavioural treatments are proven to be effective in reducing disability in CLBP patients, at this time it is unclear which treatment ingredients, or combinations between these, account for the positive findings reported. Recently, it is suggested that research should focus more on establishing which specific treatment works for whom (moderators), as well as to establish their mechanisms of change (mediators) (Kazdin and Nock, 2003; Turk, 2005; Vlaeyen and Morley, 2005). In this

line of reasoning, it has been proposed that especially fearful CLBP patients may benefit from EXP. After all, when patients do not hold any expectations of damage attached to certain activities, extinction of these associations is of no use, and there is no fear that needs to be reduced. On the other hand, treatments that do not target at reducing pain-related fear may not be suitable for highly fearful CLBP patients, since these patients will remain to be hampered by their fear. With respect to GA, highly fearful patients may not select activities as treatment goals that they fear and therefore consider to be out of reach. There are few studies that corroborate the proposition that highly fearful patients benefit more from treatments that target at reducing pain-related fear. For example, it was demonstrated that usual general practitioner care was less suitable in patients high in pain catastrophizing, whereas a minimal intervention focusing on risk factors for chronicity appeared to be more appropriate in patients holding elevated fear avoidance beliefs (Jellema et al., 2006). Another study showed that a fear avoidance based treatment, consisting of education about fear and avoidance in addition to graded exercise supervised by a physical therapist, was successful in reducing disability as compared to usual care in acute LBP patients high in fear avoidance beliefs, but to be counterproductive for those low in these beliefs (George et al., 2003). Thus, according to the suggestion to match treatments to patients characteristics, it would be expected that EXP, and not GA, should be especially successful in fearful patients, whereas smaller differences between treatments would be expected for patients with lower levels of fear.

In addition, it is worthwhile to examine the treatments' active ingredients (mediators). Theoretically, it would be expected that the specific mechanisms of change of EXP might be its systematic and explicit reduction of pain catastrophizing and perceived harmfulness of activities, and that these qualities cannot be ascribed to GA. However, some studies have shown that the effectiveness of active treatments was mediated by reductions in pain catastrophizing, even though these treatments did not explicitly aim at diminishing pain-related fear or catastrophizing cognitions (Burns et al., 2003; Spinhoven et al., 2004; Smeets et al., 2006d). Notwithstanding, even if GA would also be effective by these mechanisms of change, these are hypothesized to be more profound in EXP.

### *Verifying treatment fidelity*

An important but often neglected aspect attached to treatment outcome studies is the verification of the fidelity of the treatments delivered (e.g. the extent to which the treatments are carried out as intended) to ensure honest and genuine comparisons between the treatments of evaluation. This applies especially to complex and extensive psychological interventions, which are at higher risk of compromised treatment fidelity (Moncher and Prinz, 1991; Pereplechikova and Kazdin, 2005). In the absence of treatment fidelity checks, conclusions about treatment effectiveness can become seriously obscured as the internal as well as external validity of the study may be at stake (Moncher and Prinz, 1991; Waltz et al., 1993; Bellg et al., 2004). Although treatment fidelity may comprise various aspects, both with respect to treatment delivery by the therapists and receipt by the patients, we believe that at the minimum it should be assured whether the therapists adhered sufficiently to the treatment manuals. Besides ensuring the occurrence of valuable components in the treatment delivery, the assessment of protocol adherence also includes a verification of the non-occurrence of

protocol deviations, e.g. the absence of treatment contamination (Moncher and Prinz, 1991; Waltz et al., 1993; Borrelli et al., 2005).

## Outline of this dissertation

The general theme of this dissertation concerns the role of pain-related fear in CLBP patients, the assessment of pain-related fear, and implications of pain-related fear for clinical management of CLBP. This dissertation comprises seven studies that are organized into three divisions.

### Part II: Outlining and testing the contribution of pain-related fear in CLBP

The thesis focuses first on theoretical outlining, as well as testing, of the contribution of pain-related fear in (C)LBP. Because research on the fear avoidance model in chronic musculoskeletal complaints has been accumulating since the review by Vlaeyen and Linton (2000), in *chapter 2* a comprehensive overview is given about the current scientific evidence for the various components of this model, and about the contribution of pain-related fear during various stages of LBP. Furthermore, the implications of pain-related fear for treatment are discussed, as are future directions for research.

Since most studies available present support for the role of pain-related fear in clinical patients who are likely to differ from the average LBP patient in duration and severity of complaints, *chapter 3* presents a study examining pain-related fear in a community sample with (C)LBP. In this study, one of the propositions of the fear avoidance model is tested, namely whether fear of movement/(re)injury mediates the relationship between pain catastrophizing and functional disability, by examining several prerequisites for mediation.

### Part III: Assessment of pain-related fear

In the third part of this dissertation, various assessment methods of pain-related fear are addressed, not only with respect to CLBP patients, but to the general population without LBP complaints as well. *Chapter 4* presents a study that examines whether the TSK, for which adequate psychometric properties already have been demonstrated in CLBP patients, is suitable for application in the general population. An adapted version of the TSK (the TSK-G) is used in this study that enables people without LBP to also complete it. The factor structure of the TSK-G is examined in both people with and without LBP from the general population, as well as its construct validity by examining whether it is associated with measures of pain catastrophizing and general health status.

Because of several drawbacks attached to the original PHODA, and the fact that its psychometric properties are thus far unknown, in *chapter 5* we examine whether a shortened electronic version of this measure (the PHODA-SeV) is a psychometrically sound instrument to determine perceived harmfulness of activities in CLBP patients. The data storage of the PHODA-SeV is automatic, and its administration is less time-consuming and more standardized than the original PHODA. The factor structure, internal consistency, test-retest reliability, stability, and construct validity of the PHODA-SeV are examined, and preliminary normative data are presented.

Because of some limitations attached to self-report measures, in *chapter 6* a study is presented that examines whether it may be worthwhile to study fear of movement/(re)injury with implicit attitude measures, that is without the need for the awareness and reflection of the patient on what is being assessed. By the application of two implicit measures, it is investigated whether CLBP patients demonstrate more implicit fear of movement/(re)injury than healthy controls, and whether the two implicit measures are related to each other, and to an explicit self-report measure of fear of movement/(re)injury.

#### **Part IV: Clinical implications of pain-related fear for CLBP management**

The fourth part of this dissertation focuses on treatment implications of pain-related fear. Several small-scaled studies already demonstrated the effectiveness of EXP in highly fearful CLBP patients (Vlaeyen et al., 2001; Linton et al., 2002a; Vlaeyen et al., 2002a; Vlaeyen et al., 2002b; Boersma et al., 2004; de Jong et al., 2005b), but its effectiveness remains to be established by a randomized study with various therapists and settings, and ample patients with differing degrees of pain-related fear. In *chapter 7*, a study is presented that examines the effectiveness of EXP as compared to GA directly and six months after treatment in CLBP patients reporting at least moderate fear of movement/(re)injury, by means of a multi-centre randomized controlled trial. Furthermore, in line of matching treatments to patient characteristics, it is examined whether EXP is especially more effective than GA in highly fearful patients. In addition, this study focuses on studying specific mediating mechanisms, namely whether the effectiveness of EXP as compared to GA is mediated by decreases in pain catastrophizing and perceived harmfulness of activities, which are the presumed profound working mechanisms of EXP.

In *chapter 8*, a study is described that presents the design and application of a method to assess protocol adherence and treatment contamination (MPAC). With the use of this instrument it is verified whether treatment comparisons between EXP and GA in the randomized controlled trial described in the previous chapter are fair, in the sense that both treatments are equally and adequately carried out according to their treatment manuals. The validity and interrater reliability of this measure are determined. Furthermore, a representative sample of treatment session recordings is evaluated by two independent raters with this method. From their ratings, it is determined whether both treatments are performed as prescribed by the treatment manuals (adequate protocol adherence), and whether these treatments are not contaminated with elements that were prohibited by these manuals (absence of treatment contamination).

Finally, in *chapter 9 (part V)*, a summary of the main findings is presented together with a general discussion in which these findings are positioned within the available literature. Furthermore, the methods and procedures of the studies are critically appraised, and clinical implications as well as recommendations for future studies are outlined.





## **Part II**

**Outlining and testing the contribution  
of pain-related fear in chronic low back pain**



## **Chapter 2**

### **The fear-avoidance model of musculoskeletal pain: Current state of scientific evidence**

**Maaïke Leeuw, Mariëlle E.J.B. Goossens, Steven J. Linton  
Geert Crombez, Katja Boersma, Johan W.S. Vlaeyen**

**This chapter is published as:**

Leeuw M, Goossens MEJB, Linton SJ, Crombez G, Boersma K, Vlaeyen JWS. The fear-avoidance model of musculoskeletal pain: Current state of scientific evidence. *Journal of Behavioral Medicine* 2007;30(1):77-94.

## **ABSTRACT**

Research studies focusing on the fear-avoidance model have expanded considerably since the review by Vlaeyen and Linton (Vlaeyen J. W. S. & Linton, S. J. (2000). Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. *Pain*, 85(3), 317-332). The fear-avoidance model is a cognitive-behavioural account that explains why a minority of acute low back pain sufferers develop a chronic pain problem. This paper reviews the current state of scientific evidence for the individual components of the model: pain severity, pain catastrophizing, attention to pain, escape/avoidance behaviour, disability, disuse, and vulnerabilities. Furthermore, support for the contribution of pain-related fear in the inception of low back pain, the development of chronic low back pain from an acute episode, and the maintenance of enduring pain, will be highlighted. Finally, available evidence on recent clinical applications is provided, and unresolved issues that need further exploration are discussed.

## 1. INTRODUCTION

Despite the fact that almost anyone may experience an acute low back pain (LBP) episode at least once during his/her life, only a small minority develop a chronic low back pain (CLBP) problem (Waddell, 2004c). A traditional biomedical model that focuses upon structural and biomechanical abnormalities cannot sufficiently explain chronic pain and its associated disability. A biopsychosocial model may provide a better account because it also includes psychological and social factors (Waddell, 2004a). One psychological factor that has received much attention in the case of (chronic) pain is fear. Before describing the role of fear in the domain of pain, let us first briefly examine fear and anxiety in general as an introduction.

### **Fear and anxiety: a brief introduction**

Fear is the emotional reaction to a specific, identifiable and immediate threat, such as a dangerous animal or an injury (Rachman, 1998). Fear may protect the individual from impending danger as it instigates defensive behaviour that is associated with the fight or flight response (Cannon, 1929). The three main components of fear (interpretation of the stimulus as threatening, increased sympathetic arousal, and defensive behaviour) are loosely coupled and can change at different paces (Rachman and Hodgson, 1974; Lang, 1979). Defensive escape behaviours reduce fear levels in the short term, but may strengthen the fear in the long run. Not least, they may prevent disconfirmation of the patient's beliefs and sometimes they make the feared outcome more likely to occur. Three recent findings in fear research are notable. *First*, people with a phobia (intense irrational fear) do not necessarily have a history of being exposed to a traumatic incident (Hermans et al., 2006). *Second*, during extinction procedures, no unlearning takes place, but rather new learning that leaves the original association between the conditioned and unconditioned stimulus intact, making relapse likely (Bouton, 2002). *Third*, individual differences in vulnerabilities exist that can affect how likely fear will be experienced, acquired or maintained over time (Mineka and Zinbarg, 2006).

Anxiety, in contrast to fear, is a future-oriented affective state and the source of threat is more elusive without a clear focus. Even though the components of anxiety are similar to those of fear, they are less intense. Furthermore, whereas fear motivates the individual to engage in defensive behaviours, anxiety is associated with preventative behaviours, including avoidance. An important component of anxiety is hypervigilance, which occurs when the individual engages in environmental scanning for potential sources of threat, selectively attends to threat-related rather than neutral stimuli, broadens the attention prior to the detection of these stimuli, and narrows the attention in the presence of such stimuli (Eysenck, 1992). Both avoidance behaviour and hypervigilance reduce anxiety in the short term, but may be counterproductive in the long run. Clinically, the distinction between fear and anxiety is less evident, and these terms are frequently used interchangeably.

## 2. FEAR AND PAIN

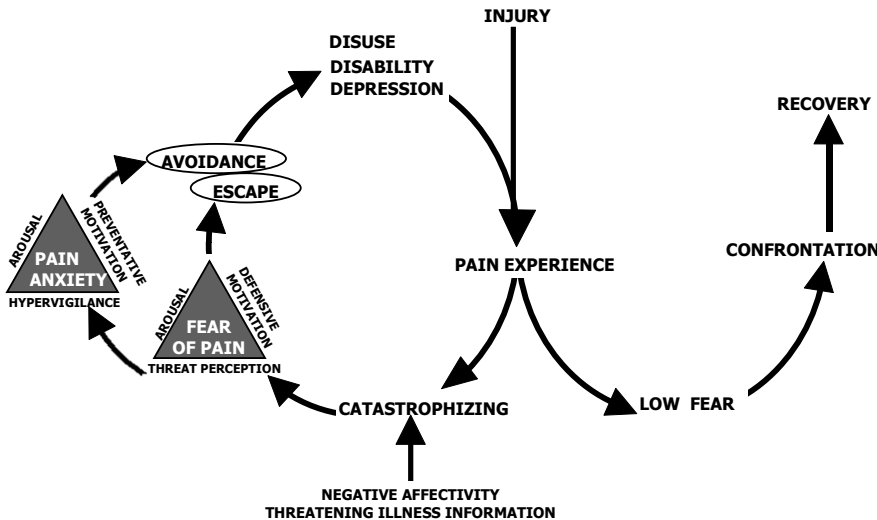
Building upon the knowledge derived from the general fear and anxiety research, the contribution of fear and anxiety in the area of (chronic) pain has been increasingly

acknowledged. Pain-related fear and anxiety can best be defined as the fear that emerges when stimuli that are related to pain are perceived as a main threat. The fear and anxiety response comprises psychophysiological (e.g. heightened muscle reactivity), behavioural (e.g. escape and avoidance behaviour), as well as cognitive (e.g. catastrophizing thoughts) elements. Depending on the current concerns, different stimuli are attended to and possibly misinterpreted. Fear of pain, fear of work-related activities, fear of movement, and fear of (re)injury have been described as often occurring in patients suffering from pain.

**The fear-avoidance model of pain**

Based upon previous work (Lethem et al., 1983; Philips, 1987; Waddell et al., 1993), Vlaeyen et al. (1995a) and Vlaeyen & Linton (2000) proposed a cognitive behavioural model of CLBP that has become known as the fear-avoidance model. This particular model provides an explanation of why CLBP problems and associated disability develop in a minority of those experiencing acute LBP. The basic tenet of the model is that the way in which pain is interpreted may lead to two different pathways (Figure 1).

**Figure 1**  
The fear-avoidance model of chronic pain. Based on the fear-avoidance model of 'Vlaeyen JWS & Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: A state of the art. *Pain* 2000;85(3):317-332' and the fear-anxiety-avoidance model of 'Asmundson GJG, Norton PJ, Vlaeyen JWS. Fear-avoidance models of chronic pain: an overview. In: Asmundson GJG, Vlaeyen JWS, Crombez G, editors. *Understanding and treating fear of pain*. Oxford: Oxford University Press, 2004. pp. 3-24'



When acute pain is perceived as non-threatening, patients are likely to maintain engagement in daily activities, through which functional recovery is promoted. In contrast, a vicious circle may be initiated when the pain is catastrophically

(mis)interpreted. These dysfunctional interpretations give rise to pain-related fear, and associated safety seeking behaviours such as avoidance/escape and hypervigilance, that can be adaptive in the acute pain stage, but paradoxically worsen the problem in the case of long-lasting pain. The long-term consequences, such as disability and disuse, in turn may lower the threshold at which subsequent pain will be experienced.

The terms fear and anxiety are often used interchangeably with regard to pain. As emphasized above, the distinction between anxiety and fear is theoretically correct, but difficult to make in a clinical context, especially in chronic pain where the threatening stimulus (pain) is constantly present. Asmundson et al. (2004) tried to differentiate between fear and anxiety and presented an updated fear-anxiety-avoidance model. In this updated model, there is no longer a direct link between fear and avoidance behaviour, because by definition one cannot avoid a threat that is already present. Instead, they added an anxiety pathway in the anticipation of pain to the fear pathway in the presence of pain. In Figure 1, both the fear and anxiety pathways are graphically presented. Since it is not yet evident that this updated model added value to the original one, and because of the remaining debate about clinical distinctiveness of fear and anxiety, we will use the term pain-related fear in the remainder of this review.

### **The fear-avoidance model of pain: evidence for its components**

Since the review by Vlaeyen & Linton (2000), numerous research studies have been undertaken in an attempt to further validate the fear-avoidance model, basically focusing on the interrelationships between its constitutive components. Below, the current scientific evidence for the separate elements of the fear-avoidance model will be discussed. The component ‘pain-related fear’ and its contribution in various stages of LBP will be separately addressed in section 3, because of its central position in the fear-avoidance model and the large body of research dedicated to it.

#### *Pain severity*

Until recently, the research literature on the fear-avoidance model did not ascribe a role of importance to pain intensity. In their review, Vlaeyen & Linton (2000) concluded that pain intensity is not a primary factor in avoidance behaviour or disability. It has even been provocatively stated that ‘pain-related fear is more disabling than pain itself’ (Crombez et al., 1999b). It is possible that this may be the case in a subgroup of chronic pain patients, but it has become clear that high pain intensity is in itself a threatening experience that drives escape and avoidance (Eccleston and Crombez, 1999). Indeed, numerous recent studies have shown that pain intensity has a considerable contribution in explaining disability. For example, pain was strongly related to functional disability during the acute stage of LBP (Sieben et al., 2005a; Gheldof et al., 2006), whereas previous LBP history and pain intensity were the best predictors of future disability (Sieben et al., 2005b). Such a relationship may be important for chronic pain as well (Mannion et al., 2001; van den Hout et al., 2001a; Woby et al., 2004a; Boersma and Linton, 2005a; Peters et al., 2005; Sorbi et al., 2006; Leeuw et al., 2007c). From these current studies, it can be concluded that the association between pain and disability both during the acute and chronic stages of pain may be more important than previously suggested.



### *Pain catastrophizing*

The term catastrophizing was first coined by Ellis (1962) for the process in which anxious patients dwell on the most extreme negative consequences conceivable. Pain catastrophizing can best be conceived as the cognitive element of the fear network, besides physiological reactivity and behavioural responses, and refers to the process during which pain is interpreted as being extremely threatening (Rosenstiel and Keefe, 1983; Crombez et al., 1998). Pain catastrophizing has consistently been associated with pain disability in pain patients (Turner et al., 2004; Peters et al., 2005; Sullivan et al., 2005a), as well as in the general population (Severeijns et al., 2004). For example, Sullivan et al. (2002) provided prospective support that initial levels of pain catastrophizing were related to subsequent activity intolerance.

In addition to its association with disability, pain catastrophizing may be related to intensified pain in various pain problems (Severeijns et al., 2001; Buer and Linton, 2002; Turner et al., 2002; Haythornthwaite et al., 2003; Turner et al., 2004; Peters et al., 2005; Sullivan et al., 2005a). However, all these studies are correlational, allowing no causal inferences. In this context, a study by Arntz & Claassens (2004) is notable. This study experimentally manipulated the meaning assigned to a painful stimulus and examined its influence on the pain experience. Healthy volunteers, who were led to believe that a cold metal bar to the neck was hot, rated it as more painful and ascribed more damaging properties to it than other healthy volunteers, who were led to believe that the same bar was cold. The damaging interpretation mediated the relationship between the experimental manipulation and the pain experienced. Prospective studies further demonstrated that initial pain catastrophizing is related to higher pain intensity in a variety of situations, e.g. during a subsequent painful procedure (Sullivan et al., 1995; Sullivan and Neish, 1999; Edwards et al., 2004; Sullivan et al., 2004; Vlaeyen et al., 2004b; Pavlin et al., 2005), after an operation (Granot and Ferber, 2005), and in the long-term adjustment to lower-limb amputation and phantom limb pain (Hanley et al., 2004).

There is also some evidence that pain catastrophizing may be considered as a precursor of pain-related fear. For example, Leeuw et al. (2007c) demonstrated that in LBP sufferers in the general population, pain catastrophizing was related to pain-related fear six months later, even after accounting for other contributing variables such as initial levels of pain-related fear. Also, Vlaeyen et al. (2004b) found that initial levels of pain catastrophizing were related to subsequent increases in fear during internal atrial cardioversion.

### *Attention to pain*

Pain that is intense, or that is perceived as threatening, demands attention and may interrupt ongoing activities (Eccleston and Crombez, 1999; Pincus and Morley, 2001). Various studies have demonstrated that excessive attention to pain is dependent upon the presence of pain-related fear (Peters et al., 2000; Keogh et al., 2001; Goubert et al., 2004c; Van Damme et al., 2006). For example, studies that found decreased cognitive task performance in fearful LBP patients interpreted this to be the result of pain-related fear directing attention to pain (Eccleston et al., 1997; Crombez et al., 1999a; Rode et al., 2001; Crombez et al., 2002a; Peters et al., 2002; de Gier et al., 2003). Furthermore, diminished attentional bias due to treatment appeared to be mainly explained by reductions in pain-related fear, suggesting that pain vigilance results

from pain-related fear (Dehghani et al., 2004). Also pain catastrophizing has been shown to increase the attentional demand of pain in experimental settings (Crombez et al., 2002a; Van Damme et al., 2002b; Van Damme et al., 2004b; Vancleef and Peters, 2006). Although it might be suggested that excessive attention to pain mediates the relationship between pain-related fear and increased pain intensities, there is only little evidence supporting this idea (Crombez et al., 2005). Instead, pain-related fear and pain vigilance seem to contribute independently to the experience of pain (Arntz and Claassens, 2004; Roelofs et al., 2004b).

Insight into particular aspects of attentional processes has substantially increased in the past few years (see for a review: Crombez et al., 2005). Evidence converges upon the idea that attentional disruption by pain-related information is not the result of an initial shift of attention to the pain stimuli, but rather stems from difficulties in *disengaging* attention from these stimuli (Asmundson et al., 2005; Crombez et al., 2005; Roelofs et al., 2005). More evidence comes from studies showing that impeded disengagement seems to be pronounced in high catastrophizers (Van Damme et al., 2004b). For example, it was demonstrated that high pain catastrophizers had more difficulty in disengaging their attention from pain cues than low catastrophizers, because of the erroneous conviction that these pain cues signaled subsequent pain targets (Van Damme et al., 2002b). Additionally, it may be that attentional engagement is facilitated by the anticipation of pain (Van Damme et al., 2004a). Interestingly, it was found that an experimentally induced attentional bias was reduced during extinction, but that it easily returned when the feared stimulus was again encountered (Van Damme et al., 2006), which is highly likely in case of chronic pain. In sum, there is evidence that attention may be an important feature in pain perception as predicted by the model.

### *Escape/Avoidance behaviour*

Avoidance refers to behaviour aimed at postponing or preventing an aversive situation from occurring (Kanfer and Philips, 1970). Although in case of chronic pain it is not possible to avoid the pain, it is possible to avoid the perceived threat, in this case the activities that are assumed to increase pain or (re)injury. Avoidance behaviour might thus be reflected in submaximal performance of activities. Vlaeyen & Linton (2000) presented an impressive overview of studies demonstrating that fearful CLBP patients performed less well on behavioural performance tasks, pointing to possible withdrawal, and thus avoidance, from these tasks. A number of more recent studies also support this assumption. For example, pain-related fear was associated with decreased speed in both preferred and fast walking (Al-Obaidi et al., 2003), with weakened muscle strength (Al-Obaidi et al., 2000; Goubert et al., 2005b), and with diminished performance on physical tasks (Geisser et al., 2000; Vowles and Gross, 2003). For instance, it was shown that in addition to experimentally induced pain anticipation, fear avoidance beliefs were related to diminished physical task performance (Pfingsten et al., 2001).

### *Disability*

Disability refers to problems executing daily life tasks and activities, in the home as well as the work situation. Disability may be a logical consequence of prolonged avoidance behaviour and hypervigilance. It seems clear that avoidance behaviours as

defined above have an impact on daily life tasks. Also hypervigilance may contribute to disability. The selective attention to pain-related stimuli, and the associated difficulty with disengaging from these stimuli, might occur at the cost of vigilance to the usual tasks of daily life. Indeed, a large body of research found that CLBP patients with heightened levels of pain-related fear report increased disability (Turner et al., 2000; van den Hout et al., 2001a; Denison et al., 2004; Staerkle et al., 2004; Woby et al., 2004a; Boersma and Linton, 2005a; Goubert et al., 2005b; Peters et al., 2005; Leeuw et al., 2007c). However, whether the association between pain-related fear and disability is mediated by both avoidance/escape behaviours and hypervigilance has not been examined yet.

### *Disuse*

In addition to the development of functional disability, frequent and enduring avoidance of daily activities may also result in gradual deterioration of a person's muscular system and fitness. The term "disuse syndrome" refers to the physiological and psychological effects of a reduced level of physical activity in daily life (Verbunt et al., 2003a). Thus far, lower or equal daily life activity levels were found when comparing CLBP patients with healthy controls, although the number of studies is rather small (Nielens and Plaghki, 2001; Verbunt et al., 2001; Spenkelink et al., 2002).

Two other aspects of disuse seem relevant: physical deconditioning, which can either be expressed in weakened muscle strength or reduced aerobic fitness, and disordered muscle coordination during physical activity (so-called guarded movements). Although diminished muscle strength was consistently found in CLBP patients as compared to healthy controls (Brox et al., 2005; Kramer et al., 2005; Verbunt et al., 2005), this might be accounted for by submaximal performance owing to pain-related fear (Al-Obaidi et al., 2000; Goubert et al., 2005b; Verbunt et al., 2005). Therefore, aerobic fitness level seems to be a more suitable measure of physical deconditioning. Generally, the physical fitness of CLBP patients is found to be either lower or equal to that of healthy subjects (Verbunt et al., 2003a). Only one current study succeeded in demonstrating that CLBP patients have lower aerobic fitness than matched healthy counterparts, though this could not be explained by pain-related fear or other relevant variables (Smeets et al., 2006c). Other studies also failed to establish significant relationships with regard to physical deconditioning as proposed by the fear-avoidance model (Wittink et al., 2000; Wittink et al., 2002; Verbunt et al., 2003b; Brox et al., 2005).

Another aspect related to disuse is disturbed trunk muscle coordination that CLBP patients can present during physical activity. One study demonstrated that experimentally induced fear of pain had only subtle effects on the erector spinae during walking, while leaving the global pattern of trunk kinematics unaffected (Lamoth et al., 2004). More evidence was found by Geisser et al. (2004), who showed that among CLBP patients pain-related fear was not only associated with reduced lumbar flexion and greater EMG in full flexion, but also to abnormalities in the muscle activity during flexion from the standing position. These changes in musculoskeletal functioning and flexion may be important for the understanding of how pain may interfere with daily life functioning.

In sum, neither lower physical activity levels nor the physical consequences of long-term avoidance behaviour in CLBP patients were unambiguously confirmed. It

may be that these patients avoid certain activities, that is those that are assumed to increase pain or (re)injury, but in general maintain their physical activity levels comparable to healthy people who may be, in general, inactive themselves (Verbunt et al., 2003a). This might explain their serious functional limitations despite unaffected levels of activity. To elucidate the actual contribution of disuse in CLBP, in the future it may be important to study the consequences of LBP *within* an affected person instead of between groups (Verbunt et al., 2003a).

### **Vulnerabilities**

An intriguing question concerns the etiology of pain catastrophizing and pain-related fear. Are there certain vulnerabilities that predispose individuals to overly attach negative appraisals to pain? If so, does this incline one to misinterpret ambiguous, but harmless, physical sensations as painful or threatening, thereby predisposing an individual to develop pain complaints (Pincus and Morley, 2001; Arntz and Claassens, 2004)? Asmundson et al. (1999) suggested that fear of pain may be secondary to the fundamental fear 'anxiety sensitivity', that is the fear of anxiety-related sensations associated with pain. Indeed, there is evidence that pain-related fear is related to anxiety sensitivity (Asmundson et al., 2000; Zvolensky et al., 2001), as well as injury/illness sensitivity (Vancleef et al., 2006), which in turn contributes to the broader construct of trait anxiety (Reiss, 1991). Along the same lines, the more general construct of negative affectivity (or neuroticism) has also been shown to lower the threshold at which pain is perceived as threatening (Goubert et al., 2004c; Gheldof et al., 2006). Because negative affectivity, anxiety sensitivity, and fear of pain seem to share some potential overlap, and at the same time differ in terms of generality, Keogh and Asmundson (2004) proposed to follow the idea put forward by Lilienfeld (1993) in the context of personality research, in approaching these constructs as part of an interrelated hierarchy. In this hierarchy, the more general negative affectivity is placed at the top, and the more specific anxiety sensitivity and fear of pain at lower levels. Although it is not quite clear which specific trait is important for the development and course of LBP, it may be suggested that certain relatively stable vulnerability traits are important and should not be overlooked.

Although speculative, it may be that individuals with an increased vulnerability to catastrophizing and pain-related fear are less changeable in their fear avoidance beliefs than those without these vulnerabilities. For clinical practice, it would therefore be important to identify these stable personality traits, since their presence may require a different, and more thorough approach, than when addressing fear avoidance beliefs in patients without these vulnerabilities.

Above, the support for the separate components of the fear-avoidance models is outlined. To date, only two studies examined the interrelationships of various elements of the fear-avoidance model simultaneously using structural equation modeling. Goubert et al. (2004c) showed that pain catastrophizing and pain-related fear mediated the relationship between neuroticism and pain vigilance, and that pain vigilance was associated with heightened pain severity. Moreover, neuroticism was found to moderate the relationship between pain severity and pain catastrophizing. Cook et al. (2006) demonstrated that pain catastrophizing was related to pain-related fear, depression and disability, that pain-related fear was related to depression and

disability, and that both depression and disability were related to pain severity. Although not allowing for causal inferences, which actually constitute the basic premises of the model, these studies do support the associations between various elements of the fear-avoidance model.

### **3. PAIN-RELATED FEAR DURING VARIOUS STAGES OF LBP**

Although the fear-avoidance model was originally developed to explain the transition from acute to chronic pain, most of the research has been carried out with chronic pain patients, underscoring the evidence for pain-related fear as the mechanism behind the maintenance of chronic pain and pain disability. Nevertheless, more recently, researchers have begun to study the influence of the role of pain-related fear in individuals with acute pain episodes, and a handful of studies have even gone a step further and focused on pain-related fear as a risk factor for the development of pain episodes in pain-free individuals.

#### **Pain-related fear as a maintaining factor of CLBP**

One of the key mechanisms in the maintenance of anxiety disorders is the prolonged use of safety behaviours including escape from, and avoidance of, threatening situations, and the selective attention towards the possible sources of threat. As reviewed above, accumulating evidence seems to underscore the contribution of escape/avoidance behaviour and hypervigilance in the area of chronic pain. Fearful pain patients may continuously scan their environment for potential signals of pain, and when the detected stimuli are interpreted as a threat, the attention is more likely to stay attached to those stimuli (Van Damme et al., 2002b; Van Damme et al., 2004a). Not only may these disrupted attentional processes be associated with increased disability because of leaving less attention available for other tasks and activities, but also by leading to intensified pain (Arntz et al., 1994). The other possible maintaining factor is avoidance/escape behaviour. When CLBP patients have detected a threatening situation, in this case a potential dangerous activity, they will either escape from the activity by not performing it at all, or avoid the activity by submaximal performance or the use of safety behaviours (Al-Obaidi et al., 2000; Goubert et al., 2005b). This behaviour fuels the pain-related fear of these patients even further, first because it prevents them from discovering that these activities are harmless, and second because it promotes disability and disuse. Escape/avoidance behaviour and disrupted attentional processes may therefore contribute to the maintenance of CLBP.

#### **Pain-related fear as a risk factor for the development of chronic LBP**

Due to differences in stringency in evaluating studies for their reviews, Linton (2000) concluded that fear avoidance beliefs were associated with both the inception and continuation of LBP, whereas Pincus et al. (2002) concluded that there was too little support for this conclusion and that instead distress and depressive mood affected the transition to chronicity. More recently, several studies have investigated pain-related fear during the acute pain phase (Gheldof et al., 2005; Gheldof et al., 2006). In acute LBP patients, pain-related fear was found to be associated with diminished participation in activities of daily life (Buer and Linton, 2002), greater perceived disability (Swinkels-Meewisse et al., 2003b; Grotle et al., 2004b; Swinkels-Meewisse

et al., 2006a; Swinkels-Meewisse et al., 2006b), greater work loss (Grotle et al., 2004b), more frequent sick leave (Gheldof et al., 2005), and diminished performance on a lifting task (Swinkels-Meewisse et al., 2006a). Due to its associations with escape/avoidance behaviour already during the acute pain phase, pain-related fear might contribute to the development of a chronic pain problem.

Several prospective studies indeed go further in demonstrating that fear avoidance beliefs can influence the transition to CLBP and associated outcomes, such as disability and sick leave. Picavet et al. (2002) showed that both heightened pain-related fear and pain catastrophizing during the acute phase increased the risk of future CLBP and disability. Accordingly, in acute LBP patients it was shown that initial elevated pain-related fear (Swinkels-Meewisse et al., 2006c), fear avoidance beliefs about work (George et al., 2005), or rising levels of pain-related fear in the beginning of a new episode (Sieben et al., 2002) were predictive of subsequent higher disability. Heightened initial levels of pain-related fear were also found to be related to decreased probability of returning to work and greater probability of being on sick leave (Fritz et al., 2001; Boersma and Linton, 2005b; Storheim et al., 2005), and to the recurrence of LBP and care seeking four years later (Burton et al., 2004). Two studies found only minor contributions of pain-related fear during the acute pain phase (Boersma and Linton, 2005a; Sieben et al., 2005a), whereas one study found pain-related fear to be unrelated to future disability (Sieben et al., 2005b). However, in the study of Sieben et al (2005b) highly fearful patients were underrepresented, and this may have resulted in an underestimation.

### **Pain-related fear as a vulnerability factor for the inception of acute LBP**

The fear-avoidance model proposes that fear avoidance beliefs may come into play when an acute pain problem emerges. It therefore seems sensible to assume that fear avoidance beliefs also exist in pain-free people. These beliefs are probably based on previous LBP experiences, personality traits (Goubert et al., 2004c), or observational learning (Mineka and Zinbarg, 2006). Recent studies indeed demonstrated fear avoidance beliefs to be present in pain-free people. For example, pain-free people in the general population reported pain-related fear comparable to, or only slightly lower, than acute and chronic pain patients (Houben et al., 2005b). Furthermore, misconceptions, such as the belief that LBP is caused by tissue damage, were very common in the general population (Goubert et al., 2004a). During pain, these misconceptions may trigger pain-related fear.

These fear avoidance beliefs may act as a vulnerability factor for the inception of a new LBP episode in pain-free people. Fearful people may be more inclined to misinterpret ambiguous physical sensations as threatening or painful, and therefore have an increased likelihood to experience pain (Pincus and Morley, 2001; Arntz and Claassens, 2004). There is indeed some evidence that fear avoidance beliefs may heighten the probability of subsequently developing a new pain episode. In currently pain-free people, pain-related fear substantially increased the risk of the subsequent inception of LBP and diminished physical functioning (Linton et al., 2000). Additionally, in workers who were currently pain free, a higher risk for future LBP onset was found for those who catastrophized about pain (Linton, 2005), and for those with heightened pain-related fear (Van Nieuwenhuyse et al., 2006). Since these studies included participants who were pain-free during the past year, it is important to

recognize that the included participants were not pain-free in the sense that they had never experienced LBP before. Due to its high prevalence (Waddell, 2004c), it is very likely that these people had suffered from previous LBP episodes, which also might have influenced the initiation of a new episode.

#### **4. PAIN-RELATED FEAR AND TREATMENT**

The aforementioned findings suggest that pain-related fear may not only be associated with the inception of a LBP episode, but also with the transition from acute to CLBP, and the maintenance of a chronic pain problem. From this knowledge we may deduce that pain-related fear could be an important target for intervention. This, however, raises the question of how pain-related fear might best be tackled within treatment.

##### **Pain-related fear: an impeding factor for treatment?**

A working alliance is of importance for positive treatment progress (Horvath and Symonds, 1991; Mead and Bower, 2000), but pain-related fear could hamper the formation of this alliance. Although rather speculative, fearful CLBP patients may be more inclined to keep searching for biomedical explanations and solutions for their pain problem, may have a history with dissatisfactory health care encounters, may strive for pain relief instead of restoration of activities, and they may more frequently have co-morbid disorders (Kerns et al., 1999; Hadjistavropoulos and Kowalyk, 2004). Because of these features, frustration may be triggered in both the patient and the therapist, thereby negatively affecting the therapeutic relationship (Hadjistavropoulos and Kowalyk, 2004). More research is required to investigate whether pain-related fear actually impedes the patient-therapist relationship.

Patients' pain-related fear may be fed by the interaction with health care providers. It might be possible that not only the health care providers' facial expressions, but also the threatening diagnostic labels they use (e.g. back injury, ruptured disc, degenerative changes) may activate the fear network (Boston and Sharpe, 2005). Because of the use of diagnostic terms that may be interpreted by patients as evidence for serious pathology, health care providers may inadvertently increase the threat value of pain. Additionally, health care providers may also hold fear avoidance beliefs themselves, although fortunately the beliefs of health care providers generally appear to be in accordance with current guidelines (Koes et al., 2001), advising the continuation of normal activities including work even when in pain, and discouraging sick leave (Linton et al., 2002b). However, fear avoidance beliefs were found to some degree in health care providers. For example, many would advise patients to avoid painful movements, some believed that sick leave is an adequate treatment, and some worried when their patients were in pain during exercise (Linton et al., 2002b). Most interestingly, these beliefs were related to treatment recommendations. The more strongly therapists held fear avoidance beliefs (Rainville et al., 2000; Houben et al., 2004; Poiraudeau et al., 2006), and the stronger the biomedical orientation of therapists (Houben et al., 2005c), the more it was recommended that CLBP patients should restrict work and physical activities. In addition to the detrimental consequences of inadequate advice, these fear avoidance beliefs of health care providers can induce or strengthen those of their patients,

especially in fearful CLBP patients who are likely to be sensitive even to minor hesitations or inconsistencies.

It is not clear whether fearful CLBP patients can benefit optimally from traditional health care. Fearful CLBP patients may sustain their avoidance behaviours regardless of the treatment message, when the disadvantageous consequences of avoidance behaviour are neither explicitly addressed, nor managed. In traditional health care, and especially in health care providers with a more biomedical orientation, the message that activity is beneficial is often accompanied by a note of caution especially if the pain increases. It is not surprising then, that fearful CLBP patients may respond with more safety and avoidance behaviours to these treatments than those who are less fearful, thereby gaining less benefit from regular treatments. One study indeed suggested that fear avoidance beliefs counteracted the beneficial effects of an exercise-based therapy (Al-Obaidi et al., 2005), and another indicated that pain catastrophizing hampered the efficacy of usual practitioner care (Jellema et al., 2006). On the contrary, it has been shown that especially for high fear avoiders a regular exercise program was more effective than usual general practitioner care (Klaber Moffett et al., 2004), whereas lower fear avoidance beliefs were associated with less favorable outcomes following a stabilization exercise program (Hicks et al., 2005). One of the questions, of course, is how these treatments have been delivered, and whether pain-related fear and the associated safety behaviours were adequately addressed.

### **Effectiveness of cognitive behavioural programs**

Some recently reported cognitive behavioural treatments have included methods to address pain-related fear. These programs have demonstrated promising results thus far, indicating that it can be beneficial to focus on decreasing pain-related fear. In line with the idea of matching treatments to specific patient characteristics (Vlaeyen and Morley, 2005), it was demonstrated that usual general practitioner care was less suitable in cases of high pain catastrophizing, whereas a minimal intervention focusing on risk factors for chronicity appeared to be more appropriate in the case of elevated fear avoidance beliefs (Jellema et al., 2006). Another study showed a fear avoidance based treatment, consisting of education about fear and avoidance in addition to graded exercise, to be successful in reducing disability as compared to usual care in acute LBP patients high in fear avoidance beliefs, but to be counterproductive for those low in these beliefs (George et al., 2003). It might therefore be that the presence of fear avoidance beliefs may debilitate outcome when usual treatment is applied, whereas fear-avoidance based treatments fail to be effective in the absence of pain-related fear.

Several cognitive behavioural programs, comprising various treatment techniques directed at changing cognitions and improving behavioural participation in activities, have currently been developed (Morley et al., 1999; Guzman et al., 2001; van Tulder et al., 2001). These studies show that the effectiveness of these programs in reducing disability was associated with decreases in pain catastrophizing (Burns et al., 2003; Sullivan and Stanish, 2003; Spinhoven et al., 2004; Sullivan et al., 2005b; Smeets et al., 2006d) and pain-related fear (Jensen et al., 2001; Mannion et al., 2001; McCracken et al., 2002; Sullivan and Stanish, 2003; Klaber Moffett et al., 2004; Woby et al., 2004b). For example, it was demonstrated that a short cognitive behavioural



intervention was successful in reducing fear avoidance beliefs, together with diminishing the risk of pain continuation, over and above the rather strong natural recovery rate for acute LBP (Linton and Ryberg, 2001). The association between pain-related fear and the effectiveness of treatments is also highlighted by brief educational sessions. As was previously demonstrated (Symonds et al., 1995; Burton et al., 1999; Moore et al., 2000), a straightforward educational campaign comprised of the advice to stay active and discouraging medical interventions, resulted in long-lasting reductions in fear avoidance beliefs in the community as well as in physicians, and in less workers compensation claims for LBP (Buchbinder et al., 2001; Buchbinder and Jolley, 2005). Taken together, these results suggest that cognitive behavioural programs, and even brief educational sessions, can effectively diminish disability, which might be due to reducing fear avoidance beliefs and pain catastrophizing.

Pain-related fear might therefore be an essential target for successful interventions. Yet, the above-mentioned treatments were not developed to specifically and systematically address pain-related fear. It can be suggested that methodically addressing and diminishing fear might be even more powerful. Exposure *in vivo*, which is a comprehensible example of an idiosyncratic treatment designed only for those with elevated levels of pain-related fear, may therefore be the most suitable treatment to achieve the most significant reduction of pain-related fear and associated disability.

### **Exposure in vivo**

Following the suggestions of Philips (1987), exposure *in vivo* treatment was developed to gradually confront patients with activities they feared and avoided for a long time due to the belief that these might be damaging for the back. This treatment consists of four components: the choice of functional goals, education about the paradoxical effects of safety behaviours, the establishment of a fear hierarchy, and graded exposure to feared activities in the form of behavioural experiments. Due to gradual and repeated encounters with these feared activities, this treatment may provide patients with the most convincing evidence that expected detrimental consequences of these feared activities are in fact a catastrophic overestimation. Due to this, their fear may diminish and functional abilities might be promoted. A detailed description of exposure *in vivo* is provided elsewhere (Vlaeyen et al., 2004a).

For the establishment of the fear hierarchy of activities, that constitutes the foundation for the rest of the therapy, the Photograph Series of Daily Activities (PHODA; Kugler et al., 1999) may be the most suitable instrument. It is comprised of photographs of various daily activities, of which patients have to indicate to what extent they think that these can be harmful to the back. The PHODA provides therapists with a tangible tool that specifically assesses the patients' beliefs about the harmfulness of certain activities for the back. Besides clinical purposes, the PHODA has also been utilized as a measurement tool, and has been shown to have excellent psychometric properties (Leeuw et al., 2007b). Several experimental single case studies, including two to six patients each, have demonstrated the effectiveness of exposure *in vivo* as compared to graded activity in fearful CLBP patients, by reporting impressive reductions in pain-related fear and disability, as well as increases in activity levels in the home situation (Vlaeyen et al., 2001; Linton et al., 2002a; Vlaeyen et al., 2002a; Vlaeyen et al., 2002b; Boersma et al., 2004; de Jong et al.,

2005b). Although the size of these improvements was less substantial for the Swedish research group (Linton et al., 2002a; Boersma et al., 2004), it does provide support for extending generalization of exposure in vivo treatment across settings and therapists. Interestingly, the educational part also produced improvements in self-reported pain-related fear and catastrophizing, although these improvements gained strength after subsequent exposure in vivo. Measures of actual behaviour, including accelerometer data, and self-reported ability to perform pre-defined goals, improved after the exposure sessions and not after the educational session (de Jong et al., 2005b). Finally, the resumption of daily activities in patients suffering from complex regional pain syndrome was associated with decreases in pain-related fear, but not with a reduction in pain-intensity, which occurred only later in time (de Jong et al., 2005a). A randomized controlled trial comparing usual graded activity with exposure in vivo in fearful patients, currently performed by the first author et al., will provide additional information on the external validity and long-term cost-effectiveness.

### **Generalization of exposure in vivo**

An important assumption behind exposure in vivo is that the repeated experiences of being able to perform various activities without pain increases or inducing (re)injury to the back during treatment will extend to activities during daily life. However, in line with findings from experimental fear research (e.g. Bouton, 2002), several experimental studies suggest that generalization of these corrective encounters is rather limited in chronic pain patients. In patients high in pain catastrophizing, exposure to a particular movement (Crombez et al., 2002b; Goubert et al., 2002), and even to a variety of movements (Goubert et al., 2005b), resulted in subsequent diminished overprediction of pain for those particular movements, but this effect failed to generalize to other dissimilar movements. It might be that these patients learn an exception to the rule ‘activities hurt’ rather than to change their general beliefs regarding the detrimental consequences of activities. Interestingly, these patients on the other hand tended to overgeneralize pain, such that once a certain movement had been painful it was assumed that it would remain painful in the future (Goubert et al., 2005a). For exposure in vivo treatment, it would be interesting to investigate whether the overprediction of injury is as difficult to generalize to dissimilar movements as the overprediction of pain.

Translated into treatment implications, these results emphasize the importance of practicing a wide variety of activities and movements, not only during treatment, but also in the home environment of the patient (Rowe and Craske, 1998a; Rowe and Craske, 1998b). Adding cognitive techniques, such as in behavioural experiments, might also be valuable because of explicitly challenging and correcting erroneous expectations. Although there is no data as yet comparing exposure in vivo to behavioural experiments, it can be suggested that the addition of behavioural experiments may optimize generalization of corrective encounters.

## **5. CONCLUSION AND UNEXPLORED ISSUES**

Summarizing, it can be stated that there is accumulating support for the fear-avoidance model. As predicted from the vast literature on fear and anxiety, pain-related fear is associated with catastrophic (mis)interpretations of pain, hypervigilance, increased

escape and avoidance behaviours, as well as with intensified pain intensity and functional disability. Disrupted attentional processes, and especially impeded attentional disengagement in the case of threatening stimuli, were found in fearful CLBP patients, as well as avoidance behaviour as reflected in submaximal physical task performance. There is less evidence for the existence of disuse in CLBP patients. Studies on aerobic fitness level, which seems the most appropriate measure of physical deconditioning, have demonstrated inconsistent findings thus far. Therefore, more research is required to establish whether CLBP patients actually have lower physical activity levels and experience associated physical consequences. Perhaps the disordered motor coordination is a more salient aspect of disuse than reduced fitness. Furthermore, evidence is accumulating that pain severity plays a more important role in disability than previously assumed. Finally, several personal vulnerabilities, such as fundamental fears or neuroticism, may influence whether someone will respond fearfully to a painful experience.

Pain-related fear may augment the vulnerability to develop new LBP episodes in currently pain-free people, may increase the risk for the continuation of LBP complaints, and may maintain complaints when they have become chronic. Some prudence is required though, since it is not straightforward that once patients respond with pain-related fear to pain, they will inevitably become mired in a vicious circle leading to the development and maintenance of enduring pain. This conception disregards the erratic and recurrent course of LBP for most people. For example, the number and duration of episodes, fluctuations in pain intensity, level of disability, occupational consequences, and health care use can vary considerably between people suffering from LBP. These individual differences in the course and consequences of LBP should be kept in mind when interpreting the contribution of pain-related fear during different stages of pain. In fact, the research on contextual factors influencing the occurrence and reinstatement of fear waits to be extended to the area of pain.

It is also important to note that the fear-avoidance model only accounts for the development and maintenance of problems in a subgroup of CLBP patients. Various other factors besides pain-related fear may determine the progress of LBP. Furthermore, supportive evidence for the interrelations between the separate constructs of the fear-avoidance model does not provide evidence for *causal* interrelationships. Causality can only be established when each construct is experimentally manipulated. It would be highly interesting to investigate the progression of interrelationships between the variables of the fear-avoidance model in a large cohort of acute back pain patients of whom some will develop subsequent chronic pain, for example by using Latent Trajectory Classes Analyses. Additionally, exploring the effects of treatments by manipulating the individual constructs in the fear-avoidance model would shed light on the actual causal contributions. Until now, a large number of cross-sectional and prospective studies at least point to the potential validity of the fear-avoidance model. For the future, a number of interesting issues remain that require further exploration. Some of these will be highlighted below.

### **Pain-related fear: when is it adaptive and when dysfunctional?**

Due to their protective and defensive function in the case of impending or immediate danger, it may be concluded that fear and anxiety are adaptive mechanisms for an individual, and in most cases they are indeed. But when do they become

dysfunctional? This is a complex issue as we lack objective measures of what is adaptive. One way to approach this question is to consider contextual issues, e.g. the presence or absence of real harm, and the consequences of fear on function and identity. In acute stages of pain, when the pain is the result of an acute injury, pain-related fear is likely to be adaptive as it will direct attention towards the injury, thereby enhancing the probability of necessary care, e.g. by withdrawal from usual activities so that healing process is facilitated (Wall, 1979). However, in the case of enduring pain, when the pain may no longer be explained by injury, fear can become dysfunctional. The persevering use of avoidance and escape behaviours in this chronic stage may be the result of misdirected problem-solving, in which patients keep seeking to solve their insoluble problem of pain release (Aldrich et al., 2000). However, in the case of chronic pain, these behaviours prevent the individual from disconfirming the level of threat. In addition, hypervigilance and avoidance behaviour will interrupt cognitive activity, interfere with daily tasks, and threaten major life goals and personal identity (Morley and Eccleston, 2004). Thus, one could assert that fear, and pain-related fear in particular, is never dysfunctional, but it is the prolonged engagement in these safety behaviours that are dysfunctional.

### **When to target pain-related fear?**

Most treatments are developed for chronic pain patients, thereby disregarding earlier phases during which the development of CLBP can be tackled. A promising area lies in preventing dysfunctional reactions when a new pain episode is initiated. Since misconceptions about LBP and its management seem very common in the general population in addition to some health care providers, these provide an ultimate opportunity for prevention. Educational campaigns have proved their value in adaptively changing LBP beliefs and associated benefits (Symonds et al., 1995; Burton et al., 1999; Moore et al., 2000; Buchbinder and Jolley, 2005). Although it can be questioned whether changing fear avoidance beliefs will lead to actual behaviour change (de Jong et al., 2005b), the potential beneficial effects of educational campaigns should be more extensively explored.

Another challenge lies in the prevention of the development of enduring pain once an acute LBP episode is established. An important example of this is the development (Linton and Hallden, 1998) of the Örebro Musculoskeletal Pain Screening Questionnaire, which identifies several psychological ‘yellow flags’ which signal a heightened risk for chronification of complaints (Linton and Boersma, 2003; Boersma and Linton, 2005b). In addition to identification of those at risk, prevention strategies should be developed that target these risk factors, thereby putting detrimental developmental processes in pain on hold (Linton and Andersson, 2000; Linton and Ryberg, 2001; Linton et al., 2005; Linton and Nordin, 2006).

### **Pain-related fear in patients with specific pain diagnoses**

So far, most of the work on pain-related fear has focused on patients with non-specific medical diagnoses such as back pain, chronic headache (Norton and Asmundson, 2004; Nash et al., 2006), fibromyalgia (de Gier et al., 2003; Turk et al., 2004), whiplash disorder (Sterling et al., 2003; Nederhand et al., 2004), and chronic fatigue syndrome (Silver et al., 2002; Nijs et al., 2004). Theoretically, however, there is every reason to believe that fear processes would be applicable to specific pain problems as

well. Indeed, the contribution of pain-related fear has been shown to extend to patients with osteoarthritis (Heuts et al., 2004), burn pain (Sgroi et al., 2005), knee injury (Kvist et al., 2005) and neuropathic pain (de Jong et al., 2005a). It will be a challenge for future researchers to integrate fear reduction techniques in medical treatment and care for these patient groups.

### **The object of fear**

Fear of pain and associated avoidance behaviours may not be the only kind of fear associated with chronic pain. Given the debilitating consequences of longstanding avoidance behaviours, patients may present with multiple fears and concerns, including the inability to work (functional), having surgery (health), having to sell the house (financial) and being a burden to family members (social; Davies, 2003). For chronic pain patients, another important concern may be social isolation that occurs in response to their diminished participation in daily life. Recently, it was shown that sensitivity to pain and sensitivity for social rejection were mutually reinforcing (Eisenberger et al., 2006). It therefore may be that concerns about social isolation in pain patients inadvertently increase their pain threshold.

Morley and Eccleston (2004) proposed to broaden the fear-avoidance model to those areas that pose a possible threat to an individual's identity, using Carver and Scheier's goal-oriented model of self-regulation (Carver and Scheier, 1998). Following this theory, a distinction is made between approach and avoidance goals. Approach goals are those that the individual is hoping for, whereas avoidance goals consist of situations that have a negative value. Fear is regarded as the emotional reaction to a movement towards an avoidance goal. Chronic pain is associated with moving towards many avoidance goals, e.g. being dependent upon others, and may therefore induce fear. Although this theoretical framework appears a promising avenue for broadening current fear-avoidance models, research efforts have just started. For example, Karoly and Ruehlman (1996) showed that in chronic pain patients, goal conflicts predicted anxiety levels over and above pain-related variables.

### **Task persistence instead of avoidance**

As mentioned before, fear is not always associated with avoidance behaviour. Rachman (2004) eloquently describes situations in which fearful individuals are willing to approach a frightening object or situation. In some cases, fear may even be associated with persistence rather than avoidance. For example, both LBP patients who were requested to perform a lifting task (van den Hout et al., 2001b), and healthy controls who were requested to immerse their hand in cold water (Severeijns et al., 2005) with the instruction "as long as possible" demonstrated *better* performance after an experimentally induced negative mood as compared to those in positive mood. One possible account for these findings is the so-called mood-as-input model, which predicts that task performance is the result of the interaction between mood and certain stop-rules. When individuals adopt an explicit or implicit "As-Many-As-Can" stop rule (they continue until satisfaction is reached about dealing with the task), negative mood will facilitate task performance, whereas positive mood will inhibit task performance. The strength of the mood-as-input model is that it provides specific predictions, including those that are based on the fear-avoidance model, but also those that are based on a novel theoretical basis that may account for findings inconsistent

with the fear-avoidance model, such as task perseverance instead of avoidance despite pain-related fear. The validity of this model in the domain of pain, however, awaits further investigation (Vlaeyen and Morley, 2004).

### **ACKNOWLEDGEMENTS**

This study was supported by Grant nr. 1436.0002 of the Netherlands Organisation for Health Research and Development (ZonMw).



## **Chapter 3**

### **Pain-related fear in low back pain: A prospective study in the general population**

**Maike Leeuw, Ruud M.A. Houben, Ruud Severeijns,  
H. Susan J. Picavet, Erik G.W. Schouten, Johan W.S. Vlaeyen**

**This chapter is published as:**

Leeuw M, Houben RMA, Severeijns R, Picavet HSJ, Schouten EGW, Vlaeyen JWS. Pain-related fear in low back pain: A prospective study in the general population. *European Journal of Pain* 2007;11(3); 256-266.



**ABSTRACT**

A cognitive behavioural account of chronic low back pain (CLBP) proposes that the relationship between pain catastrophizing and functional disability is mediated by fear of movement/(re)injury. Several clinical studies already demonstrated the contribution of pain catastrophizing and fear of movement/(re)injury in the development and maintenance of CLBP. This study included people with low back pain (LBP) in the general population, and aimed to investigate whether fear of movement/(re)injury mediated the relationship between pain catastrophizing and functional disability, by examining several prerequisites for mediation. Data from the Dutch population-based Musculoskeletal Complaints and Consequences Cohort (DMC<sub>3</sub>) study were used, including 152 people suffering from LBP who completed both a follow-up questionnaire and a baseline questionnaire six months previously. This study was unable to demonstrate that the relationship between pain catastrophizing and functional disability was mediated by fear of movement/(re)injury, since the prerequisite that pain catastrophizing and functional disability were related, was not fulfilled. However, pain catastrophizing was significantly related to fear of movement/(re)injury six months later, above and beyond other contributing variables such as fear of movement/(re)injury already present at baseline. On its turn, fear of movement/(re)injury was related to functional disability, in addition to pain intensity. Although this study leaves some indistinctness concerning the actual relationships between pain catastrophizing, fear of movement/(re)injury, and functional disability, it does provide some evidence for the contributing role of these factors in LBP in the general population.

## INTRODUCTION

Whereas low back pain (LBP) is very common, only few people develop chronic (CLBP) complaints (Klenerman et al., 1995; Waddell, 1996; Walker, 2000; Waddell, 2004c). Fear of movement/(re)injury and pain catastrophizing are important contributors in the development and maintenance of CLBP (Asmundson et al., 1999; Crombez et al., 1999b; Vlaeyen and Linton, 2000). They are interlinked with each other in cognitive behavioural models (Lethem et al., 1983; Philips, 1987; Waddell et al., 1993), assuming that a catastrophizing response to pain initiates a vicious circle in which fear of movement/(re)injury develops, and the subsequent avoidance of activities augment functional disability by means of hypervigilance, depression, and disuse. In essence, it proposes fear of movement/(re)injury as a mediator between pain catastrophizing and functional disability.

Indeed, pain catastrophizing has been linked to pain-related fear, disability, pain intensity, depression, and psychological distress (Severeijns et al., 2001; Severeijns et al., 2004). Pain catastrophizing was also related to amplified pain experiences and disability (Sullivan et al., 2001), back pain chronicity (Burton et al., 1995), and inception of LBP in pain-free workers (Linton, 2005). Fear of movement/(re)injury was more strongly related to decreased performance on behavioural tasks (Vlaeyen and Linton, 2000), disability in daily life situations (Waddell et al., 1993; Buer and Linton, 2002), and work loss (Waddell et al., 1993) than biomedical symptoms and pain severity. Fear avoidance beliefs were related to subsequent inception of LBP in pain-free people (Linton et al., 2000; Picavet et al., 2002), and course of LBP (Klenerman et al., 1995; Linton et al., 2000; Picavet et al., 2002). The effectiveness of treatments that decrease fear of movement/(re)injury and pain catastrophizing also supports the importance of these factors (Jensen et al., 2001; Vlaeyen et al., 2001; Vlaeyen et al., 2002b; Vlaeyen et al., 2002c; Boersma et al., 2004; Spinhoven et al., 2004; Woby et al., 2004b).

To our knowledge no study has yet investigated whether fear of movement/(re)injury is a mediator between pain catastrophizing and functional disability. Additionally, previous research has mainly focused on clinical patients, who probably differ from the average LBP patient in duration and severity of physical and psychological complaints. For the current study, data from the Dutch population-based Musculoskeletal Complaints and Consequences Cohort (DMC<sub>3</sub>) were used. Supplementary to studies performed with these data (Picavet et al., 2002; Severeijns et al., 2002; Picavet and Hazes, 2003; Picavet and Schouten, 2003; Picavet and Hoeymans, 2004; Severeijns et al., 2004), the main aim was to investigate in a community sample with LBP whether fear of movement/(re)injury mediates the relationship between pain catastrophizing and functional disability. The following prerequisites for mediation were examined (Baron and Kenny, 1986): (1) pain catastrophizing is related to fear of movement/(re)injury six month later; (2) fear of movement/(re)injury is related to functional disability; (3) pain catastrophizing is related to functional disability six months later without accounting for fear of movement/(re)injury, and 4) the strength of the relationship between pain catastrophizing and functional disability decreases significantly when fear of movement/(re)injury is simultaneously accounted for.

## MATERIALS AND METHODS

### Participants and procedure

All subjects participated in the DMC<sub>3</sub> study. This study was a national study about the prevalence and course of musculoskeletal complaints in the Netherlands, carried out by the National Institute of Public Health and the Environment in collaboration with Statistics Netherlands (Picavet et al., 2000; Picavet and Schouten, 2003). The study was performed during 1998 and 1999, when the Dutch population counted more than 15 million inhabitants. During the first part of the study, a full-colour booklet of questions was sent to a sample of 8000 persons randomly drawn from the population register of 1998, who were non-institutionalized, aged 25 or older, living in the Netherlands, and stratified by 10 years age groups and gender. The booklet consisted of 28 pages of questions concerning sociodemographics and health-related issues, supplemented with several standardized questionnaires, and will be further referred to as 'baseline questionnaire'. Overall, 3664 (46%) participants returned this baseline questionnaire. Those who provided informed consent for a follow-up measurement (N = 2752) received a second booklet after six months (further referred to as 'follow-up questionnaire'). Only persons between 25 and 65 years of age were included in the analyses, leaving 1581 people returning the follow-up questionnaire. Of these, 396 (25%) indicated to suffer from low back pain at the moment of completing the follow-up questionnaire. Missing value analysis for the relevant study variables indicated that for the majority of these people more than 10 percent of data on any of the measures included was missing. In case of less than 10 percent missing values, missing values were substituted with the questionnaire specific mean of that person. Excluding people with more than 10 percent missing values on any of the measures included resulted in a group of 152 people suffering from LBP, of whom the data were analyzed in the current study.

### Measures

There were two occasions of measurement, including a baseline questionnaire and a follow-up questionnaire six months later. During the baseline questionnaire, the following measurements were obtained: Musculoskeletal complaints, pain catastrophizing, functional disability, pain intensity, and fear of movement/(re)injury in the general population. During the follow-up measurement six-months afterwards, again measurements about musculoskeletal complaints, functional disability, and pain intensity were taken, while in addition information about fear of movement/(re)injury in low back pain was gathered. An overview of the baseline and follow-up questionnaires is provided in Table 2.

#### *Musculoskeletal complaints*

In both questionnaires items were included regarding the presence of musculoskeletal complaints, referring to five different anatomical areas: (1) neck, shoulders, higher part of the back; (2) elbow or wrist/ hand; (3) lower part of the back; (4) hip or knee, and (5) ankle and foot. Questions about each anatomical area were represented by one of five coloured areas in the booklet. First, the participants were asked whether they experienced pain in that specific anatomical area during the past twelve months. When complaints were experienced, questions were asked about the anatomical site, whether

or not the pain was still present, duration and severity of the pain, course of pain, determinants of the complaints, specific complaints, and some consequences of the pain. Information about health care consumption was also gathered for each specific complaint, by inquiring about visits in the past six months to the general practitioner, specialists, physical therapists, and consumption of medication. Sociodemographic variables were also queried.

### *Pain catastrophizing*

In the baseline questionnaire, the Dutch version (Crombez and Vlaeyen, 1996) of the Pain Catastrophizing Scale (PCS; Sullivan et al., 1995) was used. The PCS consists of 13 items, on which the patients are asked to indicate on a five-point scale, ranging from 0 (not at all) to 4 (always), to which degree they experienced certain thoughts or feelings during pain. The total score ranges between 0 and 52. The PCS shows good psychometric properties (Sullivan et al., 1995; Osman et al., 1997; Osman et al., 2000; Van Damme et al., 2000; Van Damme et al., 2002a).

### *Pain intensity*

Pain intensity was questioned about as part of the assessment of general health status and well being of the patients, for which in both questionnaires the RAND-36 was used. The RAND-36 is a generic health status survey that obtains patients' own assessment of their functioning, well-being and perceived general health during the past four weeks. The RAND-36 consists of 36 items divided over nine subscales, that is (1) physical functioning; (2) social functioning; (3) role restrictions due to physical problems; (4) role restrictions due to emotional problems; (5) mental health; (6) vitality; (7) pain; (8) perceived general health; and (9) perceived changes in health status. Questions can be answered by response choices ranging from two to six. Of every subscale a summation score can be calculated, with a better health status reflected in higher scores. In the analyses of the current study, only the subscale 'pain' was included. The subscale pain was recoded such that higher scores indicate higher pain intensities. The RAND-36 is found to have sufficient psychometric properties (van der Zee et al., 1993; van der Zee et al., 1996; Essink Bot et al., 1997).

### *Fear of movement/(re)injury-general population*

In the baseline questionnaire, the Tampa Scale for Kinesiophobia - General Population (TSK-G) was included to measure fear of movement/(re)injury in the general population. This is a slightly adapted version of the Tampa Scale for Kinesiophobia (TSK; Kori et al., 1990). The TSK consists of 17 statements that have to be rated on a 4-point scale ranging from 'strongly disagree' to 'strongly agree'. The total score is calculated after inversion of items 4, 8, 12, and 16, which are phrased in reversed key. The Dutch version of the TSK has been shown to be sufficiently reliable and valid (Vlaeyen et al., 1995c; Goubert et al., 2000; Goubert et al., 2004d). In the TSK-G, all items have been rephrased in such way that people without low back pain can also complete them. For example, the item 'If I were to try to overcome it, my pain would increase' was adapted to read 'If I had low back pain and I were to try to overcome it, my pain would increase'. The TSK-G has shown to consist of one single factor, excluding the reversed key items (4, 8, 12 and 16) because of too low inter-item correlations, and item 11 because of its low factor loading (Houben et al., 2005b). The

total score can be calculated by summing up the remaining 12 items. The TSK-G appeared to have sufficient internal stability and validity both for individuals with and without back pain (Houben et al., 2005b).

#### *Fear of movement/(re)injury, short version*

In the follow-up questionnaire, a short version (TSK-SV) of the TSK (Kori et al., 1990) was included. Only people who indicated to suffer from low back pain at the time of completing the follow-up questionnaire finished the TSK-SV. The original authors of the DMC<sub>3</sub> study selected eight items (items 2, 3, 4, 6, 9, 11, 13, and 14) from the full-version TSK because of shortage of space in the booklet. Selection of these specific items was realized by removing the most skewed items of the comparable TSK-G, and by including at least one item of each of the four factors that were acknowledged at that time ('harm', 'fear of (re)injury', 'importance of exercise', and 'avoidance of activities'; Vlaeyen et al., 1995b). Because the psychometric properties of the TSK-SV have never been investigated, the reliability and validity analyses of the TSK-SV are presented below.

Of 285 participants, complete data were available for the TSK-SV. The descriptive statistics for the eight TSK-SV items are shown in Table 1. Data examination showed the reversed key item (item 4) to have low inter item correlations, and was therefore precluded from further analyses. A principal component analysis with oblique rotation on the remaining seven items revealed one single factor, explaining 40.23% of the variance (Kaiser-Meyer-Olkin Measure = .78 and Bartlett's Test of Sphericity  $\chi^2 = 214.89$ ,  $df = 21$ ,  $p < .001$ ). All items showed sufficient factor loadings (ranging from .38 to .74). The internal reliability of these items was adequate (Cronbach's Alpha = .72).

The TSK-SV correlated significantly with the TSK-G ( $r = .62$ ,  $p < .001$ ), which is even more significant because of the time-gap of six months between both measurements. Furthermore, there were significant correlations between the TSK-SV and pain catastrophizing ( $r = .51$ ,  $p < .001$ ), functional disability ( $r = .56$ ,  $p < .001$ ), and pain intensity ( $r = -.52$ ,  $p < .001$ ) underscoring the validity of the TSK-SV (Table 2).

#### *Functional disability.*

The Quebec Back Pain Disability Scale (QBPDS; Kopec et al., 1995) was used to assess the severity of functional disability. The QBPDS consists of 20 items, describing activities commonly affected by back pain. Patients rate the degree of difficulty of performing these activities on a six-point self-report scale. The reliability and validity of the QBPDS are well established (Kopec et al., 1995; Schoppink et al., 1996; Fritz and Irrgang, 2001). One benefit of the QBPDS is that the items are especially developed for LBP patients from a conceptual approach to disability measurement and empirical methods of item development. In addition, the QBPDS is probably more sensitive and responsive than other disability measures (Kopec et al., 1995; Schoppink et al., 1996).

**Table 1**

Mean, standard deviation (SD), skewness, standard error (SE) of skewness, factor loadings, and inter item (spearman) correlations for TSK-SV-items and total score for people with LBP.

N = 285		Mean (SD)	Skewness (SE)	Factor loadings	Inter item correlation						
		Item			A	B	C	D	E	F	G
A	If I were to try to overcome it, my pain would increase	2.05 (1.24)	.57 (.20)	.51							
B	My body is telling me I have something dangerously wrong	1.68 (1.01)	1.08 (.20)	.74	.35***						
C	My pain would probably be relieved if I were to exercise	2.59 (1.18)	.02 (.20)	A	.09	.14					
D	My accident has put my body at risk for the rest of my life	1.41 (.81)	1.76 (.20)	.73	.21**	.43***	.10				
E	I am afraid that I might injure myself accidentally	1.61 (1.01)	1.33 (.20)	.59	.15	.30***	-.15	.32***			
F	I wouldn't have this much pain if there weren't something potentially dangerous going in my body	1.41 (.89)	2.02 (.20)	.79	.24**	.61***	.04	.53***	.33***		
G	Pain let's me know when to stop exercising so that I don't injure myself	2.99 (1.11)	-.77 (.20)	.38	.22**	.18*	-.06	.15	.15	.12	
H	It's really not safe for a person with a condition like mine to be physically active	1.57 (.93)	1.36 (.20)	.62	.24**	.24**	.14	.36***	.33***	.34***	.22**
Total score (8 items)		15.32 (4.56)	1.00 (.20)								
Total score (7 items, without item c)		12.72 (4.32)	.98 (.20)								

Reasons for exclusion: A = small inter-item correlations

\* = p < .05 (2-tailed); \*\* = p < .01 (2-tailed); \*\*\* = p < .001 (2-tailed)

**Table 2**  
Pearson correlations between questionnaires of the baseline questionnaire and follow-up questionnaire

Baseline questionnaire	A	B	C	D	E	F	G
<b>A</b> Pain catastrophizing (PCS)	1.00						
<b>B</b> Fear of movement-general population (TSK-G)	.53	1.00					
<b>C</b> Functional disability (QBPDS)	.40	.50	1.00				
<b>D</b> Pain intensity (RAND-36)	-.40	-.44	-.65	1.00			
<b>Follow-up questionnaire</b>							
<b>E</b> Fear of movement-short version (TSK-SV)	.51	.62	.49	-.46	1.00		
<b>F</b> Functional disability (QBPDS)	.40	.45	.87	-.61	.56	1.00	
<b>G</b> Pain intensity (RAND-36)	-.47	-.42	-.67	.64	-.52	-.74	1.00

All correlations were highly significant ( $p < .001$ )

**Statistical analyses**

Several regression analyses were performed to test whether the prerequisites of mediation effects according to Baron & Kenny (1986) were met. In order to investigate the first prerequisite, namely whether the degree of pain catastrophizing predicted the level of fear of movement/(re)injury six months later, a first regression analysis was applied. Fear of movement/(re)injury at follow-up was taken as the dependent variable, while the independent variables were entered in several steps into the regression equation. First, a block of background variables was entered, consisting of age, gender, and the presence of back pain at baseline. Second, fear of movement/(re)injury, functional disability, and pain intensity were entered as relevant baseline co-variables because of their probable effect on fear of movement/(re)injury. Third, baseline pain catastrophizing was entered to test its contribution to fear of movement/(re)injury at follow-up after correction for the previously entered co-variables.

The next regression analysis analyzed the second prerequisite, namely the relationship between fear of movement/(re)injury and functional disability both measured at follow-up measurement. In this case the dependent variable was functional disability at follow-up measurement, and again, the independent variables were entered into the regression equation in several steps. First, a block of background variables was entered into the regression equation, consisting of age, gender, and chronicity of back pain. Second, co-variables were entered because of their probable effect on functional disability, in this case baseline functional disability and pain catastrophizing, and pain intensity at follow-up. Finally, fear of movement/(re)injury at follow-up was entered to test its contribution to functional disability at follow-up after correction for previously entered co-variables.

The third prerequisite was tested by investigating whether pain catastrophizing was related to functional disability six months later when fear of movement/(re)injury was not accounted for. The previous regression analysis was repeated, this time excluding follow-up fear of movement/(re)injury as a co-variable.

If the above prerequisites were met, mediational analysis would investigate whether the strength of the relationship between pain catastrophizing and functional

disability would decrease when fear of movement/(re)injury was simultaneously accounted for. For this, the beta weights of pain catastrophizing in predicting follow-up functional disability with or without accounting for fear of movement/(re)injury would be compared, which were derived from the second and third regression analyses respectively. Whether the decrease of the standardized beta was significant would be tested by means of the Sobel test (Preacher and Leonardelli, 2001).

## RESULTS

### Sociodemographics

The group consisted of 152 people who indicated to suffer from LBP at follow-up measurement. There were 58 males (38%) and 94 (62%) females in this group, and the mean age was 47.31 (SD = 10.66) years. For 131 (86%) participants of this group, the back pain was chronic at follow-up measurement (with duration for more than three months). At baseline measurement, already 147 (96%) participants out of 152 revealed to suffer from low back pain, of which 126 (83%) participants had low back pain in chronic form back at that time. At follow-up measurement, this group of LBP patients indicated to consult several kinds of health care providers concerning their low back pain in the past six months: 39% consulted their general practitioner, 42% consulted a specialist (i.e. rheumatologist, rehabilitation physician, neurologist), 42% was treated by a physical therapist, and 37% used some kind of pain medication.

To investigate whether the people that were included in the study differed from those who were excluded because of their excessive missing values, the means of these groups were compared to each other by t-tests. Compared to the excluded group with more than 10 percent missing data, it appeared that the included group with sufficient data demonstrated significantly more pain catastrophizing at baseline (difference of the means = -2.98,  $t = -2.81$ ,  $p < .01$ ), pain intensity scores at both measurements (difference of the means = 13.99,  $t = 6.46$ ,  $p < .001$  for baseline and difference of the means = 10.11,  $t = 4.58$ ,  $p < .001$  for follow-up), and functional disability at follow-up (difference of the means = -8.95,  $t = -4.53$ ,  $p < .001$ ). Both groups did not differ with regard to fear of movement/(re)injury measured at baseline and follow-up (difference of the means = .05,  $t = .07$ ,  $p > .05$ , and difference of the means = -.96,  $t = -1.82$ ,  $p > .05$  respectively), and functional disability at baseline (difference of the means = -1.47,  $t = -.54$ ,  $p > .05$ ). From these findings it is clear that the included group differs from the one that was excluded, which may affect the results.

### Regression analyses

See Table 2 for the correlation coefficients between all measurements. The skewness statistics ranged between .38 and 1.07, implying that there was no serious deviation of normality. For the first prerequisite for mediation, it was analyzed whether the degree of pain catastrophizing at baseline measurement predicted the level of fear of movement/(re)injury at follow-up measurement six months later (Table 3). The variance inflation factors were acceptable (between 1.03 and 2.22), indicating that the data were not affected by multicollinearity. Several background variables were entered first ( $R^2 = .04$ ,  $p > .05$ ), followed by the co-variables fear of movement/(re)injury, functional disability, and pain intensity measured at baseline ( $R^2$  change = .47,  $p <$



.001). Finally, pain catastrophizing was entered ( $R^2$  change = .03,  $p < .01$ ). While obviously baseline fear of movement/(re)injury was the most significant predictor of fear of movement/(re)injury six months later ( $\beta = .31$ ,  $p < .001$ ), baseline pain catastrophizing was also a significant predictor for fear of movement/(re)injury at follow-up ( $\beta = .20$ ,  $p < .01$ ). In addition, gender ( $\beta = -.26$ ,  $p < .001$ ) and functional disability ( $\beta = .27$ ,  $p < .01$ ) contributed significantly to the prediction of fear of movement/(re)injury at follow-up. Thus, the first prerequisite for mediation was met, in that baseline pain catastrophizing was significantly related to fear of movement/(re)injury six months later.

**Table 3**  
Regression analyses with pain catastrophizing as a predictor for fear of movement/(re)injury at six months follow-up (first prerequisite for mediation)

Criterion variable: Fear of movement/(re)injury <sup>2</sup>						
Step	Predictor	B (SE)	$\beta$	p	$\Delta R^2$	p
1	Age	-.04 (.02)	-.11	>.05	.04	>.05
	Gender	-2.31 (.56)	-.26	<.001		
	Presence of back pain <sup>1</sup>	1.26 (1.41)	.05	>.05		
2	Fear of movement/(re)injury <sup>1</sup>	.19 (.05)	.31	<.001	.47	<.001
	Functional disability <sup>1</sup>	.07 (.02)	.27	<.01		
	Pain intensity <sup>1</sup>	.02 (.02)	.12	>.05		
3	Pain catastrophizing <sup>1</sup>	.08 (.03)	.20	<.01	.03	<.01

<sup>1</sup> = Baseline questionnaire; <sup>2</sup> = Follow-up questionnaire  
Gender: 1 = male, 2 = female

For the second prerequisite, it was analyzed whether the degree of fear of movement/(re)injury was associated with the degree of functional disability, both measured at follow-up measurement (Table 4). The variance inflation factors were again acceptable (between 1.11 – 2.28). Several background variables were entered first ( $R^2 = .13$ ,  $p < .001$ ). In the second step, relevant co-variables consisting of baseline pain catastrophizing and functional disability, and follow-up pain intensity were entered ( $R^2$  change = .68,  $p < .001$ ). Finally, fear of movement/(re)injury at follow-up was entered ( $R^2$  change = .02,  $p < .001$ ). Not surprisingly, baseline functional disability was the most significant predictor of functional disability six months later ( $\beta = .59$ ,  $p < .001$ ). Fear of movement/(re)injury was also a significant predictor ( $\beta = .20$ ,  $p < .001$ ). In addition, pain intensity ( $\beta = .26$ ,  $p < .001$ ) appeared to be a significant predictor. Finally, age ( $\beta = .07$ ,  $p < .05$ ), gender ( $\beta = .12$ ,  $p < .01$ ) and chronicity of back pain ( $\beta = -.09$ ,  $p < .05$ ) significantly predicted functional disability. The second criterion for mediation effects was therefore also fulfilled, since fear of movement/(re)injury was significantly related to functional disability.

**Table 4**

Regression analyses with fear of movement/(re)injury as a predictor of functional disability (second prerequisite for mediation)

**Criterion variable: Functional disability<sup>2</sup>**

Step	Predictor	B (SE)	$\beta$	p	$\Delta R^2$	p
1	Age	.12 (.06)	.07	<.05	.13	<.001
	Gender	4.30 (1.48)	.12	<.01		
	Chronicity of back pain <sup>2</sup>	-4.52 (1.89)	-.09	<.05		
2	Pain catastrophizing <sup>1</sup>	-.11 (.07)	-.06	>.05	.68	<.001
	Functional disability <sup>1</sup>	.60 (.05)	.59	<.001		
	Pain intensity <sup>2</sup>	.22 (.04)	.26	<.001		
3	Fear of movement/(re)injury <sup>2</sup>	.83 (.20)	.20	<.001	.02	<.001

<sup>1</sup> = Baseline questionnaire; <sup>2</sup> = Follow-up questionnaire

Gender: 1 = male, 2 = female

Chronicity of back pain: 0 = less than three months present, 1 = more than three months present

Third, it was tested whether pain catastrophizing was related to functional disability six months later when fear of movement/(re)injury was not accounted for. The second regression analyses was replicated, this time excluding fear of movement/(re)injury at follow-up (Table 5). It appeared that pain catastrophizing was not related to functional disability ( $\beta = -.01$ ,  $p > .05$ ) when fear of movement/(re)injury was not accounted for. From this, it has to be concluded that the third requirement for mediation was not satisfied.

**Table 5**

Regression analyses with baseline pain catastrophizing as a predictor of functional disability at follow-up without accounting for fear of movement/(re)injury (third prerequisite for mediation)

**Criterion variable: Functional disability<sup>2</sup>**

Step	Predictor	B (SE)	$\beta$	p	$\Delta R^2$	p
1	Age	.09 (.06)	.05	>.05	.13	<.001
	Gender	1.84 (1.42)	.05	>.05		
	Chronicity of back pain <sup>2</sup>	-3.25 (1.96)	-.06	>.05		
2	Pain catastrophizing <sup>1</sup>	-.01 (.07)	-.01	>.05	.68	<.001
	Functional disability <sup>1</sup>	.67 (.05)	.66	<.001		
	Pain intensity <sup>2</sup>	.26 (.05)	.30	<.001		

<sup>1</sup> = Baseline questionnaire; <sup>2</sup> = Follow-up questionnaire

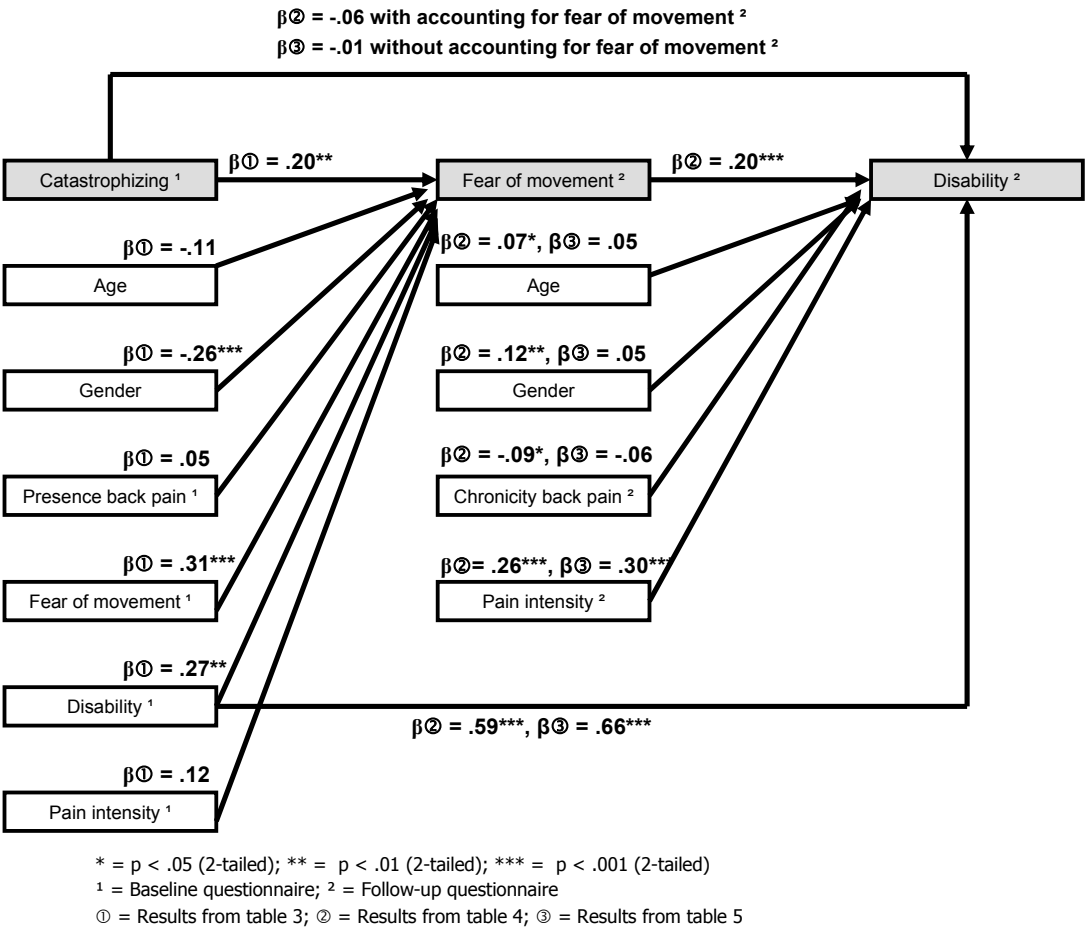
Gender: 1 = male, 2 = female

Chronicity of back pain: 0 = less than three months present, 1 = more than three months present

Fourth, because of the absence of a relationship between baseline pain catastrophizing and functional disability at follow-up, the criteria for a mediation effect were not

fulfilled, and it was therefore of no use to perform subsequent analyses towards the decrease of the association when the potential mediator was accounted for. In Figure 1 an overview of all regression analyses is depicted.

**Figure 1**  
Overview of results of regression analyses



**DISCUSSION**

The main aim was to investigate in a community sample with LBP whether fear of movement/(re)injury mediated the relationship between pain catastrophizing and functional disability, by investigating several prerequisites for mediation (Baron and Kenny, 1986). The prerequisite that pain catastrophizing was related to functional disability was not met, indicating that this study was not able to demonstrate that fear of movement/(re)injury was a mediator between pain catastrophizing and functional disability. Despite this inability to determine a mediation effect, the current study revealed other interesting findings.

First, pain catastrophizing at baseline was shown to be moderately related to fear of movement/(re)injury six months later, above and beyond the contribution of other relevant co-variables such as the level of fear of movement/(re)injury already present at baseline. This suggests that people who interpret their pain catastrophically at baseline are likely to experience increased fear of movement/(re)injury later on. This is the first study that was able to demonstrate a prospective relationship between pain catastrophizing and fear of movement/(re)injury. These results also provide strong support for the idea that the cognitive-behavioural model of fear also applies to the area of chronic pain. It also underscores the importance of pain catastrophizing as a potential vulnerability factor for the development of pain-related fear. Indeed, individuals who catastrophize about pain are likely to interpret neutral interoceptive information as 'dangerous' or 'signaling (re)injury' (e.g. Rief et al., 1998). Results also showed that men, and participants who were more functionally disabled at baseline reported increased fear of movement/(re)injury at follow-up.

In addition, fear of movement/(re)injury was cross-sectionally related to functional disability. Although the magnitude of this relationship was moderate, it suggests that people who were afraid of movement/(re)injury at follow-up also indicated to be more disabled. Pain intensity was also moderately related to functional disability. This indicates that in LBP sufferers in the general population, in addition to fear of movement/(re)injury, the level of pain determines to what degree someone will be physically impaired. This finding corroborates with other studies including LBP patients with a new episode of sick leave (van den Hout et al., 2001a), spinal pain patients with various stages of chronicity (Boersma and Linton, 2005a), and acute LBP patients (Sieben et al., 2005b) that also demonstrated pain intensity to be related to disability. Several other significant relationships were found for gender, age and chronicity of back pain, but examination of the standardized beta's revealed that they barely contributed to the dependent variable, from which it can be concluded that they were not relevant.

Unexpectedly, pain catastrophizing was not related to functional disability at follow-up. It was therefore futile to perform further analyses to test mediation. There are at least four possible explanations for the lack of an association between pain catastrophizing and subsequent functional disability. First, the study population consisted almost completely of chronic LBP sufferers. Through the stability of complaints, functional disability at follow-up was almost completely determined by baseline disability. Functional disability in the current population indeed appeared to be a stable condition, with a mean difference of 1.64 points between both measurements, which is, although significant, rather small ( $t = -2.22$ ,  $p < .05$ ). Through this, at least in this rather chronic population, other baseline variables did not have any contributing effect anymore. Studying these relationships in a rather acute LBP population will probably reveal the actual contribution of psychological factors in the transition towards chronic complaints. Second, it is possible that the association between pain catastrophizing and functional disability is confounded by a variable that was not measured in this study. Third, as can be derived from the fear avoidance model, the theoretical distance from pain catastrophizing to functional disability is larger than from fear of movement/(re)injury. The model proposes the presence of intermediate contributors, such as pain vigilance and avoidance, indicating that there is no direct relationship between pain catastrophizing or fear of movement/(re)injury and

disability. Finally, fear of movement/(re)injury was measured at the same time as functional disability, whereas pain catastrophizing was measured six months previously. In the current study, the association between pain catastrophizing and disability may have been too weak to establish statistical significance.

These findings with LBP sufferers in the general population therefore leave one with some indistinctness concerning actual theoretical relationships, since it is hard to explain why no relationship was found between pain catastrophizing and functional disability, whereas pain catastrophizing was related to subsequent fear of movement/(re)injury, and fear of movement/(re)injury on its turn to functional disability. These results are not in accordance with the theoretical cognitive behavioural account of chronic low back pain (Lethem et al., 1983; Philips, 1987; Waddell et al., 1993) that does assume a mediational relationship. Despite the fact that these results are inconclusive, it does provide support for the contribution of pain catastrophizing in subsequent fear of movement/(re)injury and of fear of movement/(re)injury in functional disability. Increasing acknowledgement of crucial psychological factors in the exacerbation of LBP complaints, in patients as well as in the general population, can provide opportunities for treatments and strategies to prevent deterioration of LBP. It should be worthwhile to correct erroneous ideas about back pain in health care providers (Linton et al., 2002b; Houben et al., 2005c) as well as in society, in order to reduce pain catastrophizing and fearful reactions to LBP. A multimedia campaign providing positive messages about LBP, advising to stay active and exercise, showed to be effective in improving fear-avoidance beliefs in the general population as well as in general practitioners (Buchbinder et al., 2001).

Several limitations of this study need to be considered. *First*, a selective non-response is possible, because on several occasions a considerable group of people did not answer certain questions, or did not fill in the questionnaire at all: for merely 152 people out of 396 suffering from LBP, complete or adequately substituted data were available. Comparison of mean scores indeed indicated some difference between the people who completed the questionnaire with little missing data and excluded people with excessive missing data. Although both groups did not differ with regard to fear of movement/(re)injury at both measurements and baseline functional disability, the included group demonstrated more baseline pain catastrophizing, higher pain intensity scores on both measurements, and more disability at follow-up. Therefore, it might be possible that the conclusions from this specific population cannot be generalized to the excluded group, nor that they can be generalized to the non-responders (Picavet and Schouten, 2003). Unfortunately, it is not known, nor possible to evaluate, why many people chose not to fill in several items or questionnaires. *Second*, it can be questioned whether the current study population was representative for LBP sufferers in the general population and differed from clinical patients, since the majority of the sample suffered from chronic complaints and a considerable proportion consulted some kind of health care. On the other hand, the majority of patients did not consult health care because of their LBP, indicating that this sample in fact did deviate from the patient population. However, in interpreting the results it should be kept in mind that a substantial subgroup might resemble a clinical population. *Third*, with this study design it was not possible to directly address the issue of causality between pain catastrophizing, fear of movement/(re)injury, and functional disability, since the findings are merely based on correlational data. *Fourth*, this study used pre-existing

data from the DMC<sub>3</sub> study (Picavet and Hazes, 2003), through which it was not possible to extend this study beyond the questionnaires used. It was possible to investigate the relationship between pain catastrophizing and fear of movement/(re)injury in a prospective manner, while the association between fear of movement/(re)injury and functional disability was studied cross-sectionally. Conclusions could be more powerful when it was also possible to prospectively investigate the association between fear of movement/(re)injury and functional disability. However, Linton (2005) emphasized the value of cross-sectional designs by demonstrating similarities between cross-sectional and prospective studies. *Fifth*, the community sample with LBP consisted of a heterogeneous group of people suffering from acute, chronic or recurrent low back pain at follow-up, of which some experienced no, acute or chronic low back pain at baseline measurement. This heterogeneity might be a disadvantage, although on the other hand the actual composition of different subgroups in the general population is maximally approached. *Finally*, all measures used in this study were self-report measures, which can be liable to various kinds of biases, such as deception and self-presentational strategies. In addition, they require the ability to consciously reflect or introspect about dysfunctional beliefs (de Jong et al., 2001; Fazio and Olson, 2003). In CLBP patients with fear of movement/(re)injury, conscious reflection or awareness of these dysfunctional beliefs can be lacking (Vlaeyen et al., 2002c).

Summarizing, this study was not able to demonstrate that in people with LBP in the general population fear of movement/(re)injury mediated the relationship between pain catastrophizing and functional disability. However, this study did demonstrate that pain catastrophizing was related to fear of movement/(re)injury six months later, and that fear of movement/(re)injury was related to functional disability.

## ACKNOWLEDGEMENTS

This study was supported by Grant nr. 1436.0002 of the Netherlands Organisation for Health Research and Development (ZonMw). We also thank Gerard van Breukelen for statistical advice.



## **Part III**

### **Assessment of pain-related fear**





## Chapter 4

### **Fear of movement/injury in the general population: Factor structure and psychometric properties of an adapted version of the Tampa Scale for Kinesiophobia**

**Ruud M.A. Houben, Maaïke Leeuw, Johan W.S. Vlaeyen,  
Liesbet Goubert, H. Susan J. Picavet**

**This chapter is published as:**

Houben RMA, Leeuw M, Vlaeyen JWS, Goubert L, and Picavet HSJ. Fear of movement/injury in the general population: Factor structure and psychometric properties of an adapted version of the Tampa Scale for Kinesiophobia. *Journal of Behavioral Medicine* 2005;28(5):415-424.

## **ABSTRACT**

In recent years, several studies have pointed out the importance of pain-related fear in the development and maintenance of chronic pain. An important instrument for measuring pain-related fear in the context of low back pain is the Tampa Scale for Kinesiophobia (TSK). Recently, a version of this questionnaire has been developed for administration among the general population (TSK-G). To determine the factor structure of the TSK-G, data from a random sample of the Dutch general population were studied separately for people who had had back complaints in the previous year, and people who had been without back complaints. For both groups the TSK-G appeared to consist of one, internally consistent, factor of 12 items. The one-factor TSK-G also appeared valid after comparison with scores on measures of catastrophizing and general health status.

## INTRODUCTION

Ever since the introduction of the term ‘kinesiophobia’ by Kori et al. (1990), scientific attention for the concept of pain-related fear in relation to chronic pain has increased. Although not everyone would go so far as to view pain-related fear as a simple phobia, a review article by Vlaeyen and Linton (2000) shows that pain researchers and clinicians alike are convinced of the importance of pain-related fear and fear-avoidance beliefs in explaining disability and the transition from acute to chronic musculoskeletal pain. One of the clinical fields of interest in this respect concerns disability as a consequence of chronic low back pain. Several studies have shown the relevance of fear-avoidance beliefs in this field (Lethem et al., 1983; Philips, 1987; Waddell et al., 1993). Tying together these findings, an etiological model based on the specific fear that physical activity will cause (re)injury has been presented (Vlaeyen et al., 1995b; Vlaeyen and Linton, 2000). This model presents a possible pathway by which injured patients either become mired in a negative spiral leading to disability, or successfully recover. A patient who catastrophizes about pain (i.e. who is convinced that his/her body is extremely vulnerable, weak and must be carefully protected from overstrain) is likely to be fearful of movement/(re)injury when experiencing pain. These fearful reactions are associated with increased muscular reactivity and avoidance, and in the long run they may lead to disuse, depression and disability. On the other hand, a patient who does not catastrophize about pain will be more likely to resume daily activities and recover successfully.

However, from a prevention point of view it would be worthwhile to find out whether fear of movement/(re)injury also occurs in the general population. If this is the case, the more fearful people might be at greater risk of developing chronic low back pain, when confronted with a low back pain episode. Research indeed showed that the presence of fear-avoidance beliefs about work in the acute state of low back pain predicted disability four weeks later (Fritz et al., 2001) and that pain catastrophizing and, especially, fear-avoidance beliefs are important in the development of low back pain, reflected in a higher risk for pain and lower physical function at one-year follow-up in previously pain-free people (Linton et al., 2000). In addition, a study of Buer and Linton (2002) showed that fear-avoidance beliefs and pain catastrophizing already existed in the general population of non-patients. Furthermore, results from the non-patients, mild pain and moderate pain groups suggested a significant contribution of fear-avoidance beliefs and pain catastrophizing in the transition from acute to chronic pain (Buer and Linton, 2002). Besides fear-avoidance beliefs in the general population, health care providers also tend to hold fear-avoidance beliefs. For example, some health care providers associate back pain with the need to avoid painful activities and view pain-reduction as a necessary requirement to return to work (Linton et al., 2002b).

Reliable and valid tools are now available to measure pain-related fear in patients, including the Tampa Scale for Kinesiophobia (TSK; Kori et al., 1990), the Fear Avoidance Beliefs Questionnaire (FABQ; Waddell et al., 1993) and the Pain Anxiety Symptoms Scale (PASS; McCracken et al., 1992). The TSK has been specifically designed to measure fear of movement/(re)injury in patients with low back pain and consists of 17 statements, which have to be rated on a 4-point scale ranging from ‘strongly disagree’ to ‘strongly agree’. Four items are phrased in reversed key.

Psychometric research carried out with the Dutch version of the TSK (Vlaeyen et al., 1995a) has shown the TSK to be sufficiently reliable ( $\alpha = .77$ ) and having predictive validity. Scores on the TSK were better predictors of disability levels in chronic low back pain patients than were, for example, pain intensity or biomedical findings (Vlaeyen et al., 1995b).

With respect to the factor structure of the TSK, Vlaeyen et al. (1995b) conducted a principal component analysis with oblique rotation on all 17 items. They found four non-orthogonal factors labeled harm, fear of (re)injury, importance of exercise and avoidance of activity. This four-factor solution explained 36.2 % of the total variance. Five items were excluded from the analysis because their factor loading was smaller than .40. Clark et al. (1996) performed a principal component analysis with Varimax rotation and found two factors, labeled activity avoidance and pathological somatic focus, explaining 49% of the total variance. All four reversed key items were excluded because item analysis showed that these items had a weak association with the total TSK score (Clark et al., 1996). Recently, Goubert et al. (2004) compared four factor models for the TSK using confirmatory factor analysis in chronic low back pain and fibromyalgia patients, namely a one-factor model (17 items), the four-factor model of Vlaeyen et al. (1995b), a one-factor model without reversed items and the two-factor model of Clark et al. (1996). The results showed that for both chronic low back pain patients and fibromyalgia patients, the two-factor model of Clark et al. (1996) provided a better fit than the four-factor model of Vlaeyen et al. (1995b), although still not adequate. Moreover, the two-factor model was found to be invariant across both patient groups, indicating that the model is robust in both pain samples. These findings were confirmed by Roelofs et al. (2004), who also provided evidence for the construct and predictive validity of the adjusted version of the TSK (Roelofs et al., 2004a).

Until now, research with the TSK has only been carried out among pain patients. As mentioned before, from a prevention point of view it would be interesting to be able to measure pain-related fear in the general population. For this purpose, the TSK has been modified (TSK-G; in which G stands for general population) to enable people without complaints to also complete it (Vlaeyen and Crombez, 1998). A first step before actually using the TSK-G in research would consist of testing it psychometrically, and comparing the results to those of the original TSK. For this purpose, the current study uses data from a large Dutch community sample that completed several questionnaires for the Dutch population-based Musculoskeletal Complaints and Consequences Cohort study (DMC<sub>3</sub>-study: Picavet et al., 2000).

The purpose of the current study was 1) to determine the factor structure of the TSK-G, 2) to determine the internal consistency of the TSK-G factors, and 3) examine the validity of the TSK-G by comparing scores on the TSK-G with data gathered in the same study on pain catastrophizing and general health status.

## METHODS

### Sample and procedure

All participants were recruited through the DMC<sub>3</sub>-study, a nationwide study of the prevalence and course of musculoskeletal complaints in the Netherlands in 1998 and 1999 (Picavet et al., 2000; Picavet and Schouten, 2003). The study consisted of two parts. First, a booklet of questionnaires was sent to a random sample of 8000 persons

aged 25 and above, living in the Netherlands, stratified by 10 years age group and sex. This sample was drawn from the Dutch population register of 1998. People who did not return the booklet received a maximum of two reminders (after three and after six weeks). In total, 3664 persons returned the questionnaire. As a part of this first booklet, people were also asked to state if they were willing to complete a follow-up questionnaire. Those who agreed received a second booklet six months after the first.

For the current study, only data from the first booklet were used. Furthermore, only persons between 25 to 65 years of age were included in the analyses. Subjects who had missing values on the TSK-G in the first booklet were excluded, leaving a total of 2240 persons with valid data. Analyses were carried out separately for people with and without back complaints at the moment of completing the questionnaire or in the year before, because the presence of back pain is assumed to influence the results. Thus, the sample consisted of 1029 persons with and 1211 persons without back complaints.

## Measurements

### *Musculoskeletal complaints*

The booklet consisted of 28 pages in full color with items regarding the presence of musculoskeletal complaints, consequences of these complaints, determinants, other physical complaints and sociodemographic variables. The musculoskeletal complaints were divided into five groups according to their location: (1) neck, shoulders, upper back; (2) elbow, wrist, hand; (3) lower back; (4) hip, knee; (5) ankle, foot. For these complaints, questions were asked about the duration, seriousness and course.

### *Catastrophizing cognitions*

The Pain catastrophizing Scale (PCS) is a 13-item questionnaire, in which patients reflect upon past experiences of pain and have to indicate on a 5-point scale (ranging from 'not at all' to 'constantly') to which degree they experienced pain-related thoughts and feelings. A higher score on the PCS points towards a higher degree of pain catastrophizing cognitions. The PCS has been found to be sufficiently reliable and valid (Sullivan et al., 1995; Severeijns et al., 2002; Van Damme et al., 2002a).

### *General health status (RAND-36)*

The RAND-36 (van der Zee and Sanderman, 1993) was included in the booklet as a measure of general health status. It is a short version of the RAND Health Insurance Study Questionnaire consisting of 36 items divided over nine subscales, being 1) physical functioning, 2) social functioning, 3) role restrictions due to physical problems, 4) role restrictions due to emotional problems, 5) mental health, 6) vitality, 7) pain, 8) perceived general health and 9) perceived changes in health status. A higher score on these subscales denotes a better health status. The RAND-36 bears close resemblance to the SF-36. Interesting for this study are the subscales 1, 2, 3, 7 and 8 (physical functioning, social functioning, role restrictions due to physical problems, pain and perceived general health), because based on previous research fear of movement/(re)injury is expected to be related to limitations in physical, but also social, functioning to general health, and possibly even to pain (Vlaeyen and Linton, 2000).

*Fear of movement/(re)injury (TSK-G)*

The Tampa Scale for Kinesiophobia (TSK) has been specifically designed to measure fear of movement/(re)injury (Kori et al., 1990) in low back pain patients. The TSK consists of 17 statements, which have to be rated on a 4-point scale ranging from 'strongly disagree' to 'strongly agree'. Four items are phrased in reversed key. The Dutch version of the TSK has been shown to be sufficiently reliable and valid (Vlaeyen et al., 1995a; Goubert et al., 2004d). For the TSK-G, items have been rephrased in such a way that people without low back pain can complete them. For example, the item 'If I were to try to overcome it, my pain would increase' was adapted to read 'If I had low back pain and I were to try to overcome it, my pain would increase'. A higher score on the TSK-G reflects a higher degree of fear of movement/(re)injury.

**Statistical analyses**

The sample consisted of persons with and without low back complaints in the previous year. Since a recent experience with low back pain could influence scores on the TSK-G, it was decided to carry out the analyses for the people with and without complaints separately. Furthermore, since a large sample was available, it was also decided to randomly split each group in half, then first carry out an exploratory factor analysis on the first half of each group, and finally test the factor structure on the second half of each group by means of a confirmatory factor analysis. Exploratory factor analysis was carried out with a Principal Axis Factor Analysis (PAF) with an Oblique rotation, using SPSS 11.0.1 (SPSS-Inc, 2001). Factors were extracted until the Eigenvalue dropped below 1 or until the Eigenvalue hardly changed between two subsequent factors, visible as a levelling off in the scree plot. Items with a factor loading below .25 were removed. If an item loaded on more than one factor, the item was removed if the difference in loading was below .1. Confirmatory factor analysis was done by means of the LISREL 8.54 statistical package (Joreskog and Sorbom, 2001). Modifications were carried out until changes in Chi-square of the modifications dropped below 4, or until the Non-Normed Fit Index (NNFI) reached the level of .90. As a goodness of fit statistic, the NNFI is uniquely used, because of its ability to estimate the goodness of fit, while accounting for large sample sizes and the frugality of a model. The first model and the final model were tested for identification (i.e. whether the estimates found are unique for the data used) by saving the fitted covariances of the respective models and using them as input. An identified model should yield a perfect fit on all parameter estimates.

Internal consistency was determined by calculating Cronbach's Alpha. Validity was studied by firstly examining relationships with other variables such as pain catastrophizing and selected subscales referring to general health status using Pearson Correlation coefficients. Furthermore, it was examined whether scores on these other variables could be predicted from scores on the TSK-G, when controlling for demographic variables, by means of regression analyses. Because of the considerable size of the sample, all statistical tests were carried out with a .01 significance level.

## RESULTS

### Sociodemographics

As described before, a total of 2240 persons were included in the analyses. Mean age was 45.2 years (SD = 11.3) and 46.3 % was male. The groups with and without low back complaints in the year prior to completing the questionnaire consisted of 1029 and 1211 persons respectively.

### Data examination

Before starting with the exploratory factor analysis both samples were randomly split in half. For the exploratory factor analysis a sample of 515 persons with low back complaints and 605 persons without low back complaints was available. Table 1 shows the descriptive statistics for the 17 TSK-G items for each of the samples. Before factor analysis, inter item correlations were computed. These showed that the four reversed key items (4, 8, 12, and 16) had generally low intercorrelations with the other TSK-G items. Furthermore, preliminary factor analyses with all 17 TSK-G-items showed that these four items were never included in factors with other items, but always formed separate factors by themselves. When looking at the content of these items, this is unexpected. The same pattern was found in psychometric research with the original TSK (Clark et al., 1996; Goubert et al., 2004d), and since feedback from clinical experience has also shown that patients find these reversed key items confusing, it was decided not to include them in any further analysis.

### Exploratory factor analysis

PAF with oblique rotation was carried out on the remaining 13 items for the samples with and without low back complaints separately. For the sample with back complaints the Kaiser-Meyer-Olkin Measure (.825) and Bartlett's Test of Sphericity ( $\chi^2 = 1345.3$ ;  $p = .000$ ) both justified continuation of the analysis. Examination of Eigenvalues in combination with the Scree plot pointed towards a one-factor solution. Subsequent factor analysis showed that item 11 had an insufficient factor loading. The remaining twelve items made up one factor. Factor loadings and reasons for exclusion for these items are shown in Table 1. Total variance explained was 31.0 %. Cronbach's Alpha for this factor was .79, which is sufficient.

Factor analysis of the sample without low back complaints followed the same pattern. The Kaiser-Meyer-Olkin Measure (.815) and Bartlett's Test of Sphericity ( $\chi^2 = 1424.3$ ;  $p = .000$ ) were again both sufficient. Also, Eigen values and Scree plot suggested the extraction of one factor, and again item 11 was removed because of not meeting the minimum factor loading criterion. Total variance explained was 29.6 %, and Cronbach's Alpha was .78 in this case<sup>1</sup>. Factor loadings and reasons for exclusion are again shown in Table 1.

---

<sup>1</sup> Cronbach's Alpha of the 17-item TSK-G was .75 for the sample with low back complaints and .72 for the sample without low back complaints.



**Table 1**  
Descriptives (mean, standard deviation (SD), factor loadings and reasons for exclusion) for TSK-G-items for people with and without back complaints

Nr	Item	With back complaints (N = 517)		Without back complaints (N = 609)	
		Mean (SD)	Factor loadings	Mean (SD)	Factor loadings
1	I'm afraid sometimes that I might injure my back if I exercise	1.9 (1.1)	.541	1.6 (1.0)	.388
2	If I had low back pain and I were to try to overcome it, my pain would increase	2.1 (1.2)	.450	2.0 (1.1)	.343
3	Back pain means that there is something dangerously wrong with your body	1.8 (1.1)	.607	1.9 (1.1)	.638
4	Back pain decreases when a person stays physically active	2.8 (1.1)	A	2.8 (1.0)	A
5	People aren't taking my medical condition seriously enough	1.7 (1.0)	.403	1.5 (.9)	.368
6	If I had long-term low back pain, the rest of my life would become endangered	1.7 (1.0)	.409	1.8 (1.0)	.542
7	Back pain means the body is injured	2.3 (1.1)	.436	2.3 (1.1)	.489
8	If back pain increases through physical activity, that doesn't mean that it is dangerous	2.9 (1.1)	A	2.9 (1.1)	A
9	I am afraid that I might injury myself accidentally	1.7 (1.1)	.632	1.7 (1.0)	.536
10	The safest way to prevent back pain from worsening, is being careful not to make any unnecessary movements	2.7 (1.2)	.296	2.6 (1.2)	.315
11	There would perhaps be less back pain if there weren't something wrong with the back	3.1 (1.1)	B	3.1 (1.1)	B
12	If I had back pain, I would try to stay physically active	3.4 (.8)	A	3.3 (.9)	A
13	Back pain means a person should stop exercising to prevent injury	1.7 (.9)	.528	1.9 (1.0)	.585
14	For a person with back complaints it is not advisable to be physically active	1.5 (.8)	.537	1.7 (.9)	.549
15	I can't do all the things normal people do, because I think I can easily get back complaints	1.6 (1.0)	.554	1.3 (.7)	.500
16	Even though something would cause me a lot of back pain, I don't immediately think it is dangerous	2.8 (1.1)	A	2.9 (1.1)	A
17	I should not have to exercise if I would have back pain	1.6 (.9)	.490	1.6 (.9)	.424
Total score (17 items)		33.6 (7.8)		32.9 (7.3)	
Total score (12 items)		22.4 (6.7)		21.7 (6.3)	

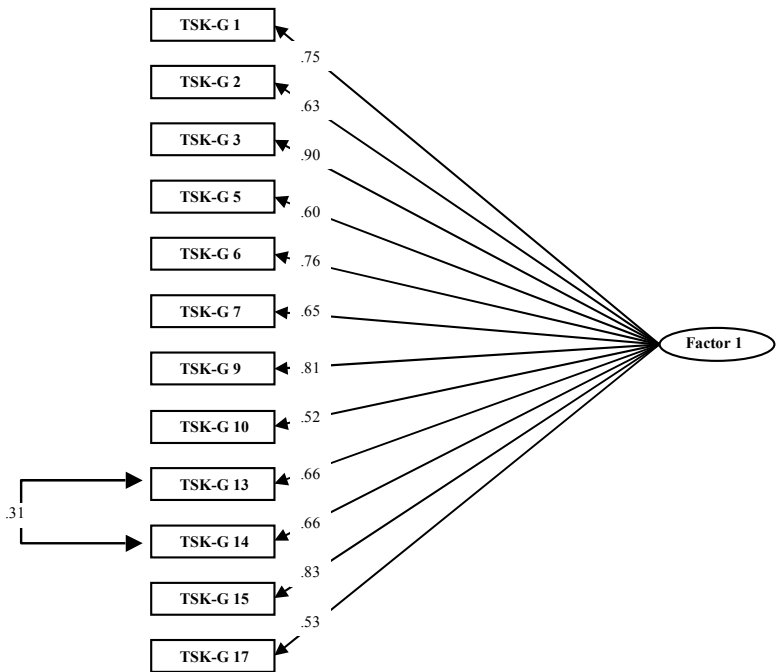
Reasons for exclusion: A = small inter-item correlations, B = factor loading below .25

### Confirmatory factor analysis

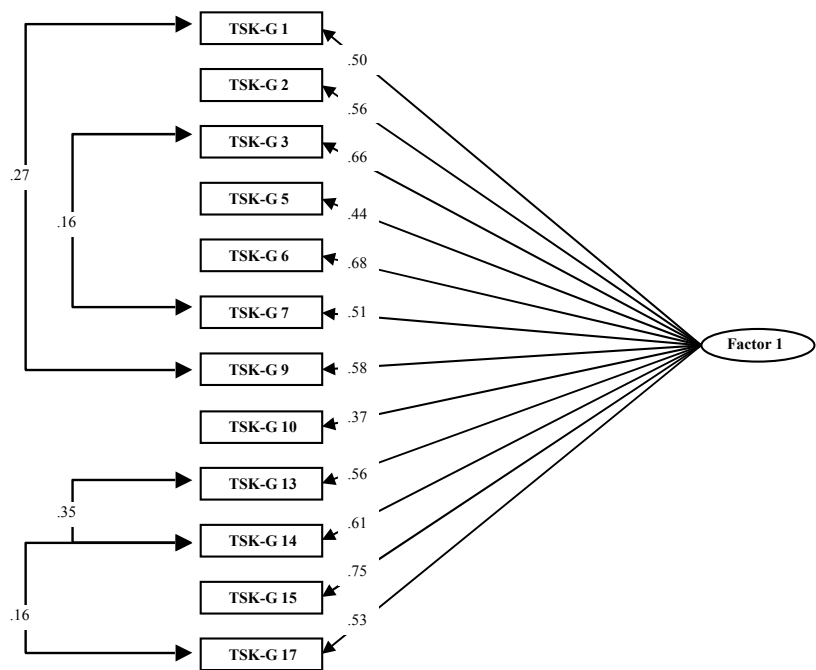
A confirmatory factor analysis was carried out using the LISREL 8.54 statistical software. Since the TSK-G-items are scored on a 4-point scale, and item scores are therefore not normally distributed, polychoric correlations were used as input instead of Pearson correlations. The one factor model that was described above was tested for the second half of the people with low back complaints (N = 514) and without low back complaints (N = 606). For the sample with low back complaints, results provided by LISREL showed that the one factor model fitted reasonably to the data (NNFI = .88). However, there was still room for improvement suggested by LISREL, through adding error covariances between certain items. After adding an error covariance between items 13 and 14, the NNFI reached the level of .91. Chi-square for this model was 161.3 (df = 53;  $p < .01$ ), and the check for identification for the first and final model showed that the model was identified. The final model is shown in Figure 1.

Next, the same procedure was followed for the second half of the sample without low back complaints. The initial fit of the one factor model was somewhat less for this sample (NNFI = .79). However, LISREL provided several suggestions for improvement through adding error covariances. The NNFI reached the level of .91 after adding four covariances between items 1 and 9, 3 and 7, 13 and 14, and 14 and 17. Chi-square for the final model was 130.4 (df = 50;  $p < .01$ ). Both the first and final model was identified. The final model is shown in Figure 2.

**Figure 1**  
Final LISREL model for people with low back complaints (factor loadings and error covariances between items)



**Figure 2**  
Final LISREL model for people without low back complaints (factor loadings and error covariances between items)



**Validity**

To determine the validity of the TSK-G, sumscores were compared to scores on measures of pain catastrophizing and selected subscales of general health status. The selected subscales from the RAND-36 were physical functioning, social functioning, role restrictions due to physical problems, pain and perceived general health. As a measure of pain catastrophizing a sumscore on the PCS was calculated. Since both samples with and without low back pain showed the same factor structure, it was decided to join both samples and carry out the following analyses on the group as a whole. However, in the regression analyses having low back pain or not was added as a covariate. Other covariates in the regression analyses were gender and age.

Pearson correlations are shown in Table 2. As can be seen, a higher degree of pain-related fear is associated with a higher degree of pain catastrophizing and lower general health status. Coefficients vary between .205 and .356 in absolute terms. Regression analyses show that TSK-G-scores are a significant predictor in all cases and even the strongest predictor when it comes to pain catastrophizing, physical and social functioning and perceived general health. Furthermore, whether or not a person has had back pain in the last year was also a significant predictor for general health status, but was hardly predictive of pain catastrophizing. Finally, it appears that being male was associated with less pain catastrophizing and a better general health status. In general however, as can be concluded from the magnitude of the adjusted R-squares, relatively little of the variance on the dependent variables was explained by the regression models (Table 3).

**Table 2**

Pearson correlations between TSK-G-scores, pain catastrophizing, and general health status

	<b>N</b>	<b>TSK-G</b>
Pain catastrophizing	2065	.356
Physical functioning	2165	-.266
Social functioning	2223	-.225
Role restrictions due to physical problem	2178	-.205
Pain	2221	-.219
Perceived general health	2176	-.287

All correlations were highly significant ( $p < .001$ )**Table 3**

Regression analyses with TSK-G-scores as a predictor for pain catastrophizing and subscales of general health status

	Pain catastro- phizing	Physical functioning	Social functioning	Role restrictions due to physical problem	Pain	Perceived general health
Adjusted R-square	.148	.181	.078	.102	.165	.148
Standardized Beta's:						
TSK-G	.375**	-.258**	-.225**	-.208**	-.215**	-.270**
Low back pain	.046	-.233**	-.144**	-.209**	-.309**	-.215**
Gender	-.144**	.146**	.082**	.126**	.132**	.060*
Age	-.013	-.198**	-.030	-.035	-.066**	-.137**

\*  $p < .01$ ; \*\*  $p < .001$ 

## DISCUSSION

The purpose of the current study was (1) to determine the factor structure of the TSK-G, (2) to determine the internal consistency of the TSK-G factors, and (3) examine the validity of the TSK-G by comparing scores on the TSK-G with data gathered in the same study on pain catastrophizing and general health status. Similar to previous psychometric research with the TSK, it appears that the four reversed key items of the TSK-G provide problems for persons completing the questionnaire. It was therefore decided to remove these four items from further analysis. Since this is a consistent finding in studies with the TSK, it might be worthwhile to rephrase these items in the same direction as the other items, and study whether they contribute to the reliability and validity of the TSK. The content of these four items does seem relevant upon the first glance.

From the remaining 13 items of the TSK-G, only one was excluded during exploratory factor analysis because of a factor loading that was too low. This item read 'There would perhaps be less back pain if there weren't something wrong with the back', and the corresponding original TSK-item was 'I wouldn't have this much pain if there weren't something potentially dangerous going on in my body'. In the adapted

item from the TSK-G, less emphasis is put on the potential threat of something being wrong with the back. Thus, it might be that this item taps less into the concept of pain-related fear, and because of this is excluded from factor analysis. Furthermore, the TSK-G item contains a double negation, which can be confusing for the person completing the scale. Perhaps a better adaptation for the TSK-G would read 'There is back pain, because there is something potentially dangerous wrong with the back'. Summing up, exploratory factor analysis showed that the TSK-G consists of one factor, made up out of 12 items, and this structure was the same for people with and without low back complaints in the year before completing the questionnaire. Internal consistency (Cronbach's Alpha) was .79 for the people with and .78 for the people without low back complaints.

Confirmatory factor analysis on the second half of the sample with and without low back complaints showed similar results to the exploratory analysis. For the persons with low back complaints an adequately fitting model was found with the addition of one error covariance. For the people without low back complaints, four error covariances were added before the model fitted sufficiently. The addition of several covariances might be an indication that the factor structure needs adaptation, since it shows that there is variance unexplained by the proposed model. This is especially necessary if these covariances show a distinct pattern. However, no pattern was observed in this case.

The validity of the TSK-G was determined by comparing scores on this measure with scores for pain catastrophizing and scores for the subscales physical functioning, social functioning, role restrictions due to a physical problem, pain and perceived general health from the RAND-36. Pearson correlation coefficients showed that a higher degree of pain-related fear was associated with a higher degree of catastrophizing cognitions and poorer health as measured on the subscales of the RAND-36. The magnitude of the correlations is not really high, but scores are not compared to a gold standard but to measures of concepts that are expected to be related. Furthermore, the correlation with the PCS is the highest, and this is also the measure that, from a biopsychosocial model, is conceptually the closest (Vlaeyen and Linton, 2000).

Regression analyses also showed TSK-G scores to be consistently associated with scores on pain catastrophizing and general health as measured by the RAND-36 subscales. It was the strongest predictor with regard to pain catastrophizing, physical and social functioning and perceived general health, when controlling for age, gender and whether or not a person had had back complaints during the last year. However, the regression models explained only a relatively small amount of the variance on the dependent variables.

Until now, only a few studies with the TSK-G have been carried out. Picavet et al. (2002) used the data from the DMC<sub>3</sub>-study to determine whether future back complaints could be predicted from scores on the TSK-G. All 17 items were used to calculate a sumscore, which was predictive of low back pain and disability measured with the follow-up questionnaire of the DMC<sub>3</sub>-study (Picavet et al., 2002). Peters et al. also used a sumscore of all 17 TSK-G items in an experiment on the predictive value of pain-related fear with regard to somatosensory hypervigilance in chronic low back pain patients. Patients completed the regular TSK, whereas healthy controls completed the TSK-G. As expected, scores of the healthy controls were found to be in between

scores of low fearful and highly fearful patients (Peters et al., 2002). The problem with both these studies is that at the time the psychometric properties of the TSK-G had not been studied yet. Although not very likely, it is possible that the results would have been different if the conclusions of the current study had been known.

Several studies have shown low back pain to be medical, social and economical problem (Waddell, 1998a). The largest part of this problem is accounted for by a relatively small group of people with chronic complaints (Nachemson, 1992). Attention of clinicians has therefore shifted towards prevention of chronicity. The TSK-G can be useful in determining people at risk for developing chronic back complaints. The current study gives preliminary data with regard to the TSK-G. More extensive data is needed on the predictive value of the (one factor; 12 item) TSK-G for developing (chronic) low back complaints. Looking at the TSK-G scores from the current study, it seems that they are a few points lower than those of acute (Sieben et al., 2002; Swinkels-Meewisse et al., 2003a), as well as chronic (Vlaeyen et al., 1995a; Vlaeyen et al., 1995b; Crombez et al., 1999b; Peters et al., 2002), low back pain patients. Further research is needed to determine which people might be at risk for developing chronic low back pain problems, but it could be assumed that cut-off points should probably also be a few points lower than those based on the regular TSK. Scores on the TSK-G were normally distributed, were similar for people with and without low back complaints and varied almost along the full range of the scale. Cut-off points on the 12-item TSK-G (with 17-item TSK-G scores between parentheses) based on the 2240 people included in the current study would be 21 (32) points for a median-split. For a division based on lower and upper tertile the cut-off points would be 18 (29) and 24 (36) points respectively. Another method for determining cut-off points on the TSK-G is the receiver operating characteristic (ROC) method (Hanley and McNiell, 1982). For the use of this method prospective data are needed on TSK-G scores and whether or not people have developed chronic low back pain. By making a graph of 'true positives' versus 'false positives' for several cut-off points, the optimal cut-off point can be determined. This method could not be used in the current study, since no data were available on the development of chronic low back pain.

Additionally, it would be interesting to determine whether on the regular TSK and the TSK-G can be related to each other. It might be expected that this is the case, since the TSK-G is merely an adaptation of the items from the TSK so the scale can be administered to people without complaints. However, it needs to be studied first whether scores gathered on the TSK-G are still useful as a measure of pain-related fear when a person has developed low back complaints, or if the regular TSK should then be administered.

In conclusion, the factor structure of the TSK-G appears to be similar for people with and without low back complaints. The TSK-G consists of one internally consistent factor, made up of 12 items. Similar to the studies by Goubert et al. (2004) and Clark et al. (1996), it is recommended to use the TSK-G without the reversed key items, or to rephrase these items in the same direction as all the other items. The other item that was excluded during factor analysis might still be used in the questionnaire, if it is phrased differently. However, both adaptations will require the TSK-G to be psychometrically studied again. Preliminary examination of the validity also points in the expected direction. High TSK scores predict pain catastrophizing, pain intensity,

and pain-related health indices, even when controlling for the presence of low back pain. It is recommended to use TSK-G as a measure of pain-related fear in general population studies.

#### **ACKNOWLEDGEMENTS**

The authors thank the National Institute for Public Health and the Environment for kindly providing the data for this study. This study was supported by Grant No. 904-59-108 of the Council for Medical and Health Research of the Netherlands Organization for Scientific Research (MW-NWO), and by Grant No. 5214 of the National Fund for Public Mental Health.

## **Chapter 5**

### **Measuring perceived harmfulness of physical activities in patients with chronic low back pain: The Photograph Series of Daily Activities - Short electronic Version**

**Maaïke Leeuw, Mariëlle E.J.B. Goossens, Gerard J.P. van Breukelen,  
Katja Boersma, Johan W.S. Vlaeyen**

**This chapter is published as:**

Leeuw M, Goossens MEJB, van Breukelen GJP, Boersma K, Vlaeyen JWS. Measuring perceived harmfulness of physical activities in patients with chronic low back pain: The Photograph Series of Daily Activities – Short electronic Version. The Journal of Pain 2007; 8(11):840-849.



**ABSTRACT**

Cognitive-behavioural models of chronic low back pain (CLBP) predict that dysfunctional assumptions about the harmfulness of activities may maintain pain-related fear and disability levels. The Photograph Series of Daily Activities (PHODA) is an instrument to determine the perceived harmfulness of daily activities in patients with CLBP. This study examined the psychometric properties of a short electronic version of the PHODA (PHODA-SeV). The results show that the PHODA-SeV measures a single factor and has a high internal consistency. The test-retest reliability and stability of the PHODA-SeV over a two-week time interval are good, with discrepancies between two measurements over 20 points suggesting true change. The construct validity is supported by the finding that both self-reported pain severity and fear of movement/(re)injury were uniquely related to the PHODA-SeV. Validity is further corroborated by the finding that patients who have received exposure in vivo, that aimed to systematically reduce the perceived harmfulness of activities, had significantly lower PHODA-SeV scores after treatment than patients receiving graded activity that did not address these assumptions. The findings support the PHODA-SeV as a valid and reliable measure of the perceived harmfulness of activities in patients with CLBP. Preliminary normative data of the PHODA-SeV are presented.

**Perspective**

This article describes a pictorial measurement tool (PHODA-SeV) for the assessment of the perceived harmfulness of activities in patients with chronic low back pain. The PHODA-SeV has good psychometric properties and can be used to elaborate on the contribution of beliefs about harmful consequences of activities to pain and disability.

## INTRODUCTION

Many people will experience acute low back pain (LBP) during their lifetime, of whom fortunately only few will develop chronic low back pain (CLBP) (Walker, 2000; Waddell, 2004c). There are several cognitive behavioural models that try to explain the development and maintenance of CLBP in a subgroup of patients by assigning a central role to the concept of pain-related fear (Lethem et al., 1983; Philips, 1987; Waddell et al., 1993; Vlaeyen et al., 1995b; Vlaeyen and Linton, 2000). Pain-related fear has been linked to the development of a new LBP episode in currently pain-free people (Linton et al., 2000; Picavet et al., 2002; Van Nieuwenhuysse et al., 2006; Leeuw et al., 2007a), and it may be a risk factor for the development of chronic complaints out of an acute episode (Picavet et al., 2002; Burton et al., 2004; Swinkels-Meewisse et al., 2006c; Leeuw et al., 2007a). Not only is pain-related fear associated with disability during the chronic stage (Geisser et al., 2000; Al-Obaidi et al., 2003; Goubert et al., 2005b; Leeuw et al., 2007a), but the treatments that aim to reduce pain-related fear in these patients have also been found to be effective (Vlaeyen et al., 2002b; Boersma et al., 2004; de Jong et al., 2005b; Leeuw et al., 2007a).

Pain-related fear is a general, overarching construct, encompassing various specific fears. For example, a more specific fear is 'fear of movement/(re)injury', which signifies the fear that certain movements may be harmful to the back by causing (re)injury (Vlaeyen and Linton, 2000). The Tampa Scale for Kinesiophobia (TSK; Miller et al., 1991) is a brief questionnaire that measures the extent to which patients with chronic pain experience fear of movement/(re)injury. Although psychometric studies have supported the reliability and validity of the TSK (Goubert et al., 2004d; Roelofs et al., 2004a), a limitation is that it does not provide information about which specific movements or activities a patient fears or avoids. For this objective, the Photograph Series of Daily Activities (PHODA; Kugler et al., 1999) may be more appropriate. The PHODA is an instrument that includes photographs of various daily activities. Back pain patients have to indicate to what extent they perceive these daily activities to be harmful to the back. Hence, the PHODA is a straightforward tool that specifically focuses on a patient's judgements about the harmful consequences of certain movements. In addition to being used to guide treatment decisions, the PHODA has been used as a measurement instrument in several studies that examined the effectiveness of exposure in vivo in CLBP (Vlaeyen et al., 2001; de Jong et al., 2002; Linton et al., 2002a; Vlaeyen et al., 2002a; Vlaeyen et al., 2002b; Vlaeyen et al., 2002c; Boersma et al., 2004; Vlaeyen et al., 2004a; de Jong et al., 2005b). Despite these previous applications, the psychometric properties of the PHODA have never been studied.

Despite the PHODA holding promise, its administration is not completely standardized, and its application can be rather time-consuming, both because it consists of 100 photographs and because there is no automatic data storage. This study aims to develop and then report on the psychometric properties of a shortened electronic version derived from the original PHODA (PHODA-SeV; both versions of the PHODA can be ordered by email: [PHODA@HSZuyd.NL](mailto:PHODA@HSZuyd.NL)). More specifically, this study aims to investigate the factor structure, internal consistency, test-retest

reliability, stability, and construct validity of the PHODA-SeV. Furthermore, preliminary normative data about the PHODA-SeV are presented.

## MATERIALS AND METHODS

### Participants and procedure

All participants were derived via their referral for participation, or actual participation, in a randomized clinical trial (RCT) that studied the effectiveness of exposure in vivo as compared to graded activity in patients with CLBP (ISRCTN88087718). The institutional ethics committee approved the research protocol, and all participants provided informed consent. Different subgroups of this study sample were included for the various analyses of this study.

#### *Screening group*

In total, 113 patients with CLBP, who were referred by physicians from various rehabilitation facilities in the Netherlands, completed the PHODA-SeV as part of the initial screening procedure for participation in the RCT.

#### *Inclusion group*

Of the patients performing this screening procedure, 85 satisfied the selection criteria, and were included in the RCT. The inclusion criteria were: (1) to have back pain for at least three months, (2) to have back pain that was not caused by a serious spinal injury (Waddell, 1998b), (3) to be between 18 and 65 years old, (4) and to report at least some fear of movement/(re)injury (TSK > 33). Exclusion criteria were: (1) to have too little disability (Roland Disability Questionnaire < 4), (2) to be illiterate, (3) to be pregnant, (4) to report substance abuse, (5) to be involved in any litigation concerning disability income, (6) to have specific medical disorders or cardiovascular diseases preventing participation in physical exercise, and (7) to present with serious psychopathology (as determined with Dutch norms of the SCL-90). The included participants then completed two measurements prior to start of treatment with a test-retest interval of two weeks, during which they also completed the PHODA-SeV. Because PHODA-SeV data of the second pre-treatment measurement were lost for one participant due to computer failure, 84 participants with both pre-treatment measurements remained.

#### *Exposure in vivo group and graded activity group*

The included patients (N = 85) were randomized, with pre-stratification on the basis of the degree of pain catastrophizing and disability, either to exposure in vivo (N = 42) or to graded activity (N = 43). The PHODA-SeV was again administered at the end of treatment. Due to refusal to complete post-treatment measurement (N = 2 out of the 10 who prematurely terminated exposure in vivo, and N = 8 out of the 14 who prematurely terminated graded activity), pre- and post-treatment measurements of the PHODA-SeV were available for N = 40 in the exposure in vivo group and for N = 35 in the graded activity group.

## Measures

### *PHODA-SeV*

The original PHODA (Kugler et al., 1999) was developed out of a need for a diagnostic tool to determine the perceived harmfulness of different physical activities and movements. The point of departure in this development consisted of eight possible movements (lifting, bending, turning, reaching, falling, intermittent load, unexpected movement, long-lasting load in stance or sit with limited dynamics), which were derived from basic movements (extending, inflecting, rotating, lateral inflecting, compression, and traction) and two manners of moving (static and dynamic). These eight possible movements were then set against four areas of daily occupations (activities of daily living, housekeeping, work, and sport and leisure time) and converted into recognizable and frequent activities instead of in terms of their biomechanics. The resulting list of movements and activities was tested, corrected, and supplemented by several experts on CLBP (human movement scientists, physical therapists, and psychologists). Finally, this resulted in 100 photographs of daily activities.

In this study, a shortened electronic version of the PHODA (the PHODA-SeV) was developed, consisting of 40 selected pictures (numbers of the original PHODA: 2, 3, 4, 7, 8, 11, 14, 15, 18, 20, 22, 23, 26, 27, 28, 29, 33, 36, 40, 44, 47, 49, 50, 51, 57, 59, 60, 61, 73, 74, 83, 85, 92, 93, 94, 95, 96, 98, 99, 100). For every basic movement category, activities were selected with variable degrees of rated harmfulness (unpublished data). A visual presentation of the PHODA-SeV is depicted in Figure 1. The patient, who sits in front of a computer monitor, is exposed to the photographs, and is requested to drag each photograph along a 'harmfulness thermometer' ranging from 0 ('not harmful at all') to 100 ('extremely harmful'). The following verbal instruction is given before the start of the test: "Please observe each photograph carefully, and try to imagine yourself performing the same movement. To what extent do you feel that this movement would be harmful to your back?" The test administrator is absent during completion, so the patient independently completes the test. Each photograph is given a rating according to its position on the thermometer. A mean total score ranging from 0 to 100 is calculated as the sum of each rating divided by 40. Due to the electronic administration, this assessment tool is highly standardized and the data are automatically stored into an electronic database. The basic properties of the original PHODA (such as the instruction, the possibility for the patient to move the pictures along the harmfulness thermometer at any time, the fact that all pictures remain visible along the fear thermometer) are maintained, except for the interaction with the therapist. The time to complete the PHODA-SeV is approximately ten minutes.

**Figure 1**  
Visual representation of the PHODA-SeV



- ❶ The photographs of the PHODA-SeV are presented in small format in a row. The patient can select each photograph by clicking on it.
- ❷ The small photograph that is selected by the patient emerges in large format.
- ❸ The selected photograph also appears in small format at this position, from which it can be dragged with the mouse to the corresponding value on the thermometer. By this means, all photographs remain visible along the thermometer. The patient can reposition each photograph at any time, by selecting it with the mouse and dragging it somewhere else.

*Tampa Scale for Kinesiophobia*

The Dutch version (Vlaeyen et al., 1995c) of the Tampa Scale for Kinesiophobia (TSK; Miller et al., 1991) was administered as a self-report measure of fear of movement/(re)injury. The TSK consists of 17 statements that can be rated on a 4-point scale ranging from 'strongly disagree' to 'strongly agree'. The total score is calculated after inversion of items 4, 8, 12, and 16, which are phrased in reversed key. The Dutch version of the TSK has been shown to have adequate psychometric properties. Several studies demonstrated the TSK to have an invariant two-factor structure among various pain diagnoses (Clark et al., 1996; Geisser et al., 2000; Swinkels-Meewisse et al., 2003b; Goubert et al., 2004d; Heuts et al., 2004; Roelofs et al., 2004a; Roelofs et al., 2007). Validity was supported by associations between the TSK and other measures of catastrophizing, fear, and disability (Vlaeyen et al., 1995b; Roelofs et al., 2004a), and with diminished behavioural task performance (Vlaeyen et al., 1995a; Crombez et al., 1999b; Geisser et al., 2000). Because several studies showed the reversed key items to be weakly related to the rest of the items (Vlaeyen et al., 1995b; Clark et al., 1996;

Goubert et al., 2004d; Roelofs et al., 2004a; Houben et al., 2005b), these were excluded from these analyses.

### *Pain Catastrophizing Scale*

The Dutch version (Crombez G and Vlaeyen JWS: The Pain Catastrophizing Scale Unpublished authorized Dutch/Flemish translation, 1996) of the Pain Catastrophizing Scale (PCS; Sullivan et al., 1995) was used to determine the degree of catastrophizing about the pain experience. Pain catastrophizing is defined as an exaggerated negative orientation towards actual or anticipated pain experiences (Sullivan et al., 1995). The PCS consists of 13 items, on which the patients are asked to indicate on a five-point scale, ranging from 0 (not at all) to 4 (always), to which degree they experience certain thoughts or feelings during pain. The total score ranges between 0 and 52. The PCS has been shown to consist of three, highly internally consistent, subscales (Sullivan et al., 1995; Osman et al., 2000; Van Damme et al., 2002a), and to have good test-retest reliability (Crombez et al., 1998). Validity of the PCS was supported by the finding that people who score high on catastrophizing report more negative pain-related thoughts, more emotional distress, and more pain when undergoing a painful procedure than those low on catastrophizing (Sullivan et al., 1995). Moreover, the PCS shows a positive association with another measure of catastrophizing (Vlaeyen et al., 1990).

### *Roland Morris Disability Questionnaire*

The Dutch version of the Roland Morris Disability Questionnaire (RDQ; Roland and Morris, 1983; Beurskens et al., 1996; Gommans et al., 1997) was used to determine the level of functional disability. The RDQ aims to measure the severity of perceived disability by questioning about limitations concerning 24 activities of daily living that can be answered by either yes or no. The RDQ has a high test-retest reliability (Roland and Morris, 1983), and is responsive to change (Grotle et al., 2004a; Rocchi et al., 2005). Validity of the RDQ was supported by its associations with other self-report and performance based disability measures (Reneman et al., 2002; Rocchi et al., 2005).

### *Current pain intensity*

Participants were asked to indicate their current pain intensity on a 100 mm visual analogue scale ranging from 'no pain at all' to 'unbearable pain', which was part of the McGill Pain Questionnaire (MPQ; Melzack, 1975).

## **Statistical analyses**

The screening group (N = 113) was used to analyze the factor structure of the PHODA-SeV. We performed an exploratory factor analysis with Oblimin Rotation. The number of factors embedded in the PHODA-SeV was determined by using the scree test and Kaiser-Meyer-Olkin criterion. For each factor found, we created a scale, and then computed a Cronbach's Alpha for that scale to assess its internal consistency.

To test the reliability of the PHODA-SeV over time, the inclusion group (N = 84) was used. The test-retest reliability was determined by calculating a Pearson correlation over the two pre-treatment measurements. Also, the stability of the PHODA-SeV was studied by testing the equality of the mean scores of both

measurements by the paired t-test, and checking equality of the variances at both time points. We furthermore plotted the difference between the two measurements against their mean, as suggested by Bland and Altman (1996). Using the average and standard deviation of the intra-individual difference between the two measurements, we computed the limits of agreement and a confidence interval for those limits (Bland and Altman, 1986).

The construct validity of the PHODA-SeV was investigated in two ways. First, we determined with Pearson correlation coefficients whether the PHODA-SeV was associated with a measure of fear of movement/(re)injury (TSK), pain catastrophizing (PCS), functional disability (RDQ), and current pain intensity (MPQ current pain intensity). For these analyses, we used the inclusion group ( $N = 84$ ). Furthermore, to adjust for common variance between these constructs, a multiple linear regression analysis was performed with the PHODA-SeV as the dependent variable and the TSK, PCS, RDQ, and MPQ current pain intensity as independent variables.

The second method we used to determine the construct validity of the PHODA-SeV was to investigate whether patients who received exposure in vivo treatment demonstrated lower post-treatment PHODA-SeV scores than those who received graded activity. Exposure in vivo treatment is designed specifically to address and diminish the perceived harmfulness of activities, as opposed to graded activity, which is designed to address the reinforcing consequences of pain behaviour (Fordyce, 1976; Fordyce, 2001b; Fordyce, 2001a). Therefore differential decreases in perceived harmfulness of activities are to be expected between these groups (for a detailed description of rationales and mechanisms of exposure treatment and graded activity, see Vlaeyen et al., 2004a and Sanders, 2002 respectively).

Complete case analyses were employed, including patients who either completed or prematurely terminated treatment and who performed the post-treatment measurement. The exposure in vivo group ( $N = 40$ ) as well as the graded activity group ( $N = 35$ ) were therefore included in these analyses. The second baseline measurement of the PHODA-SeV was used as pre-treatment measure. We verified first whether those who dropped out were comparable to those who completed treatment by univariate comparisons for pre-test PHODA-SeV, treatment group, age, gender and education. When dropouts did not differ from completers on any of these variables, we performed subsequent analyses. We performed an ANCOVA with post-treatment PHODA-SeV score as the dependent variable, group as between subject variable, and pre-treatment score included as the covariate, as this method has more power than ANOVA of the post-test or ANOVA of change (van Breukelen, 2006).

To establish norms for the PHODA-SeV, data of the screening group ( $N = 113$ ) were included in the analyses. First, we tested dependence of the PHODA-SeV score on various background variables, by means of a multiple linear regression analysis with the PHODA-SeV as the dependent variable, and age, gender, duration of complaints, and educational level as independent variables. If none of these variables was related to the PHODA-SeV score, norms were based on the score distribution in the total screening group. If dependence of the PHODA-SeV score on some background variable was found, this was taken into account by adjusting the score for the relevant background variable. A normality check was performed to test whether the PHODA-SeV scores were normally distributed. In case of a normal distribution, it is sufficient to present the mean score and standard deviation of the group, to which

individual scores can be compared by computing the corresponding z-score. In case of a non-normal distribution, deciles of the PHODA-SeV score distribution need to be presented.

## RESULTS

### Sociodemographics

The characteristics of the various participant groups are displayed in Table 1.

Data about education, work status, and complaints were unavailable for nine participants of the screening group. Valid percentages are therefore displayed where applicable.

### Factor structure and internal consistency

The item statistics are displayed in Table 2. A normality check for each of the 40 PHODA-SeV items showed no substantive deviations (apart from positive skewness for photograph 94 which had the lowest mean). Kaiser-Meyer-Olkin measure of sampling adequacy was .93, suggesting that factor analysis was appropriate. Even though the eigenvalues of the first seven factors exceeded 1 (20.47, 2.31, 1.83, 1.44, 1.25, 1.16, 1.09 respectively), the large drop in eigenvalue between factor 1 and 2 (18.16) in contrast to only small changes between the subsequent factors (max. 0.48), as well as the point of inflection of the curve of the screeplot, clearly indicate a one-factor structure. This single factor structure is furthermore supported by high factor loadings ranging between .42 and .83. This single factor explained 51.16% of the variance. The internal consistency of the total score on the PHODA-SeV, as indicated by Cronbach's Alpha, was .98. The corrected item-total correlations were high (ranging between .42 and .82), indicating that each item was moderately to highly related to the other items.

### Test-retest reliability and stability

The distribution of the PHODA-SeV scores showed no serious deviation from normality at either pre-test measurement (skewness, respectively, -.75 and -.82). The mean duration between both measurements was 14.49 (SD = 2.08) days. The mean PHODA-SeV score on the first and second measurement occasion were respectively 53.14 (SD = 19.07) and 53.57 (SD = 19.10), which did not significantly differ from each other ( $t = -.47$ ,  $p = .64$ ). As the numbers above show, the standard deviation was also very similar at both measurements. The Pearson correlation between both PHODA-SeV measurements was .90 ( $p < .001$ ). Since the standard deviation of the intra-individual difference between both measurements was 8.60, and the mean of this difference was -.44 which did not significantly deviate from zero, the limits of agreement were  $\pm 17.20$  points. The 95% confidence interval for the lower bound of the limits of agreement was -19.85 to -14.55, and for the upper bound was 14.55 to 19.85. Therefore, conservatively stated, the limits of agreement lay between  $\pm 20$  points, indicating that a change that exceeds 20 points can be regarded as a real change instead of just natural variation in scores, at least assuming a retest time interval of two weeks.



**Table 1**  
A summary of sample statistics.

	Screening group (N = 113) <sup>a</sup>	Inclusion group (N = 84) <sup>b</sup>	Exposure in vivo group (N = 40) <sup>c</sup>	Graded activity group (N = 35) <sup>c</sup>
Age (y)	43.40 ± 10.71	45.24 ± 9.45	46.05 ± 9.27	44.46 ± 8.17
Gender (% male)	45.13	51.19	47.50	60.00
Education (%)				
Low	44.10	42.20	45.00	47.10
Middle	50.00	45.80	47.50	38.20
High	5.90	12.00	7.50	14.70
Mean duration of LBP (y)	10.47 ± 8.97	11.17 ± 8.99	11.67 ± 10.06	12.11 ± 9.56
Pain radiation to the legs (%)	91.30	97.60	97.50	100
Use of medication (%)	70.20	72.60	75.00	68.60
Work status (%) <sup>1</sup>				
Work/study	40.40	38.10	35.00	40.00
Housekeeping	15.40	15.50	12.50	22.90
Unemployed/retired	7.70	9.50	12.50	5.70
Sick leave (duration in y)	26.90 (.82 ± .68)	28.60 (.88 ± .70)	27.50 (.49 ± .36)	25.70 (1.37 ± .83)
Disability pension (duration in y)	30.80 (8.12 ± 7.30)	27.40 (8.82 ± 7.14)	35.00 (9.69 ± 7.99)	25.70 (7.42 ± 5.71)
PHODA-Sev	46.20 ± 21.34	53.57 ± 19.10	52.16 ± 20.59	55.00 ± 18.35
TSK (including reversed items)	41.34 ± 7.66	41.86 ± 6.93	41.83 ± 8.45	41.40 ± 5.65
PCS	22.08 ± 12.12	23.19 ± 11.15	22.95 ± 11.50	22.83 ± 11.39
RDQ	14.50 ± 3.96	14.61 ± 3.93	15.08 ± 4.18	14.00 ± 3.99

Means ± standard deviations, or percentages, are presented

<sup>a</sup> Included for analyses of factor structure and internal consistency, and normative data

<sup>b</sup> Included for analyses of test-retest reliability and stability, and construct validity with questionnaires

<sup>c</sup> Included for analyses of construct validity by comparing treatments

<sup>1</sup> The categories of work status were not mutually exclusive, due to which more than one category could apply for one participant

**Table 2**

Item descriptives of the PHODA-SeV organized on the basis of their mean perceived harmfulness in descending order (N = 113)

No. <sup>1</sup>	Description of item	Mean (SD)	Skewness	Item total correlation	Factor loading
2	Shovelling soil with bent back	75.64 (29.77)	-1.33	.76	.78
98	Falling backwards on the grass	71.60 (31.48)	-.94	.81	.83
4	Lifting flowerpot with slightly bent back	71.42 (28.43)	-.97	.75	.77
20	Lifting beer crate out of car with slightly bent back	67.68 (28.95)	-.91	.76	.78
99	Mowing the lawn manually	62.24 (31.49)	-.51	.82	.84
29	Vacuum cleaning under coffee table with bent back	61.67 (31.71)	-.40	.80	.82
100	Drilling a hole in a stone wall above the head	56.32 (31.84)	-.29	.80	.82
57	Making the bed with bent back	55.62 (30.30)	-.27	.79	.81
47	Taking a box filled with bottles from a shelf above the head	55.57 (31.07)	-.25	.67	.69
83	Lifting a toddler (1-2 yr) from its cot with bent back	55.53 (30.76)	-.34	.79	.81
33	Mopping floor with a squeegee with slightly bent back	55.40 (30.74)	-.24	.77	.78
18	Lifting a filled basket while walking up the stairs	55.21 (28.94)	-.19	.79	.81
22	Carrying a shopping bag with one hand while walking	54.58 (30.83)	-.28	.72	.75
85	Carrying a child (5 yr) on the hip	53.62 (31.91)	-.21	.78	.79
26	Carrying rubbish bag with one hand while walking	51.16 (29.57)	-.03	.75	.77
50	Rope skipping	50.54 (34.14)	.02	.69	.71
7	Picking up shoes from floor with bent back	50.00 (32.92)	-.00	.67	.69
44	Back muscle exercise bending forward on a fitness device	49.10 (32.58)	-.05	.71	.73
49	Trampoline jumping	48.51 (34.48)	.13	.67	.68
73	Cleaning the windows with arm stretched above the head	46.52 (29.13)	.03	.76	.78
27	Clearing out the dishwasher with bent back	45.46 (29.18)	.09	.72	.74
93	Running through the forest	45.41 (32.66)	.23	.70	.72
40	Back twist exercise on a fitness device	45.36 (32.22)	.07	.73	.74
51	Abdominal muscle exercises on the floor with fitness device	43.77 (32.86)	.28	.68	.70
28	Taking a box from the sink cupboard above the head	42.67 (30.69)	.30	.71	.73
59	Getting out of bed by first placing one foot on the ground	41.66 (28.06)	.24	.67	.69
36	Leg stretch exercise on a fitness device	40.73 (31.16)	.28	.67	.69
23	Carrying two shopping bags with both hands while walking	39.09 (28.68)	.41	.64	.66
95	Cycling from a low kerb	37.39 (31.85)	.48	.69	.71
3	Lifting flowerpot squatting down	37.13 (32.11)	.62	.55	.56
11	Taking book from shelf behind oneself (with twisted back)	35.92 (29.80)	.51	.64	.66
8	Picking up shoes from floor squatting down	35.38 (28.76)	.55	.57	.58
14	Ironing in standing position	34.77 (27.34)	.67	.68	.70
92	Doing the dishes in standing position	33.35 (28.26)	.71	.64	.66
96	Looking aside while cycling	30.00 (29.88)	.92	.68	.69
60	Walking up the stairs	25.35 (25.29)	1.10	.57	.58
74	Riding a bicycle in a street with speed bumps	23.56 (25.56)	1.36	.57	.59
61	Walking down the stairs	23.11 (22.07)	1.34	.55	.57
15	Ironing in sitting position	22.65 (23.96)	1.24	.58	.59
94	Walking through the forest	17.23 (22.51)	1.77	.42	.43

<sup>1</sup> The number of the photograph refers to that of the original PHODA

Construct validity

Construct validity was first examined by investigating correlations between the PHODA-SeV and several self-report questionnaires (Table 3). The PHODA-SeV appeared to be significantly related to the TSK, PCS, RDQ, and MPQ current pain intensity. Additional evidence supporting the validity of the PHODA-SeV is that the correlation coefficients with the criterion measures were not too large (i.e., they were not .80 or greater), suggesting that the PHODA-SeV assessed something related to, but also distinct from, the other measures, especially the TSK.

**Table 3**  
Pearson correlations between the PHODA-SeV and self-report measures of fear of movement/(re)injury, pain catastrophizing, functional disability, and current pain intensity (N = 84)

	PHODA-SeV
Fear of movement/(re)injury (TSK)	.37**
Pain catastrophizing (PCS)	.23*
Functional disability (RDQ)	.30**
Current pain intensity (MPQ)	.39***

\* p < .05; \*\* p < .01; \*\*\* p < .001

To disentangle which of these constructs is uniquely related to the measure of perceived harmfulness of activities while correcting for shared variance, a multiple linear regression was performed (Table 4). Variance inflation factors were small (< 1.50), suggesting that there was no problem of collinearity. It was shown that the TSK and MPQ current pain intensity were the only unique contributors to the PHODA-SeV ( $R^2 = .25$ ). These results indicate that the perceived harmfulness of physical activities is associated with the level of fear of movement/(re)injury as well as the current level of pain severity.

**Table 4**  
Multiple linear regression of the PHODA-SeV on the self-report measures of fear of movement/(re)injury, pain catastrophizing, functional disability, and current pain intensity (N = 84)

Predictor	B (SE)	$\beta$	p
Fear of movement/(re)injury (TSK)	.96 (.39)	.29	<.05
Pain catastrophizing (PCS)	-.04 (.22)	-.02	.87
Functional disability (RDQ)	.63 (.57)	.12	.27
Current pain intensity (MPQ)	.26 (.10)	.29	<.01

$R^2 = .25$

Second, we investigated the construct validity of the PHODA-SeV by comparing post-treatment PHODA-SeV scores of those who received graded activity treatment with those who received exposure in vivo. Since 26 of 85 patients dropped out during

treatment, 10 of whom refused to complete post-treatment measurement, we first tested whether the dropouts were comparable to those who completed treatment. Univariate comparisons revealed that dropouts did not differ between groups ( $\chi^2 = .16$ ,  $p = .69$ ), and that dropouts did not differ from those completing treatment in gender ( $\chi^2 = .07$ ,  $p = .80$ ), education ( $\chi^2 = .75$ ,  $p = .69$ ), and age ( $t = .98$ ,  $p = .33$ ). As expected given the randomized treatment assignment, the two treatment groups were similar on the pre-test PHODA-SeV score (mean = 55.00, SD = 18.35 for the graded activity group, and mean = 52.16, SD = 20.59 for the exposure in vivo group;  $t = .63$ ,  $p = .53$ ) as well as other relevant measures (TSK, PCS, RDQ, age, gender, duration of complaints, education: all  $p > 0.25$ ). Post-treatment PHODA-SeV scores were 42.07 (SD = 21.07) for the graded activity group and 20.97 (SD = 17.65) for the exposure in vivo group. ANCOVA demonstrated this effect to be significant ( $p < .001$ ) and substantial ( $d = 1.13$ ). Thus, there was a large difference between both groups in post-treatment PHODA-SeV scores, with those who received exposure in vivo demonstrating lower scores than those who received graded activity. Furthermore, it was found that in the exposure in vivo condition 25 out of 40 patients (63%) demonstrated a pre- to post-test change in PHODA-SeV score beyond the limits of agreement, whereas this was only the case for 12 out of 35 patients (34%) in the graded activity condition. Note, however, that the pre-post time interval was larger than the two weeks interval from which the limits of agreement were obtained.

### Normative data

Regression analysis showed that the PHODA-SeV score was not related to any of the predictors considered (i.e. gender, age, duration of complaints, educational level: all  $p$ -values  $> .28$ ). It was therefore reasonable to calculate the PHODA-SeV descriptives for the total group, rather than to determine norms with multiple regression on demographic and other relevant patient variables (Van Breukelen and Vlaeyen, 2005). The Shapiro-Wilk normality test indicated that the PHODA-SeV scores were not normally distributed ( $p < .05$ ). Therefore, norms were based on the deciles of the PHODA-SeV distribution in the screening group instead of z-scores (see Table 5).

### DISCUSSION

The aim of this study was to investigate the psychometric properties of a short electronic version of the PHODA as a measure of the perceived harmfulness of daily activities in patients with CLBP. Advantages of this PHODA-SeV over the original PHODA are its standardized administration, the fact that it is less time-consuming, and its automatic data storage. The PHODA-SeV measures a single factor and has a high internal consistency. The test-retest reliability of the PHODA-SeV over a two-week time-interval is excellent. Limits of agreement lies within 20 points, which is satisfactory when considering the 0-100 scoring range. This implies that when one wants to detect actual changes, the change should exceed the natural variation of 20 points before it can be concluded that a real change has occurred. However, it is important to recognize the rather short time interval (2 weeks) over which the stability was determined. Over longer time intervals, the natural variation may be expected to be larger.

**Table 5**  
Deciles of the total score on the PHODA-SeV (N = 113)

Decile	PHODA-SeV score <sup>1</sup>
10	13.49
20	23.00
30	37.82
40	43.25
50	48.93
60	53.34
70	59.28
80	64.76
90	72.90
100	90.75
Mean sum score	46.20 (SD 21.34)

<sup>1</sup> The PHODA-SeV score is computed by dividing the sum of all 40 items of the PHODA-SeV by 40 and thus ranges from 0 to 100, just like each item does.

The construct validity of the PHODA-SeV is supported by consistent relationships with related constructs, such as self-report measures of fear of movement/(re)injury, pain catastrophizing, functional disability, and current pain intensity. More specifically, after correction for the common variance between these constructs, it appeared that the PHODA-SeV is specifically related to the degree of fear of movement/(re)injury as well as to pain intensity. This indicates that the more that activities are rated as harmful, the higher the level of fear of movement/(re)injury, and vice versa. Furthermore, this indicates that people experiencing more pain may be more worried that certain activities may aggravate their already painful, and potentially damaged, back than people experiencing less pain. Another explanation may be that these patients, despite the careful instruction of the aim of the PHODA-SeV, placed the photographs along the thermometer according to the expected pain that the activities would induce, which might have been primed by their current pain levels. Construct validity was further corroborated by the finding that patients in the exposure in vivo treatment condition, aimed at systematically reducing the perceived harmfulness of physical activities, had much lower post-treatment PHODA-SeV scores than the patients in the graded activity condition, during which the perceived harmfulness of physical activities was not challenged.

There are some limitations that need to be considered in interpreting these results. *First*, for the factor analyses the observation to record ratio (N = 113 with 40 observations per participant) may have been suboptimal. However, this may not pose a problem especially because of an excellent Kaiser-Meyer-Olkin measure. Also, since too small a group would complicate the factor pattern, the straightforward pattern that was found supports a single factor structure. Furthermore, simple factor structures with high loadings, as is the case with the PHODA-SeV, require smaller sample size for sufficient power than complex structures with moderate loadings. To confirm the

single factor structure of the PHODA-SeV, it will need to be cross-validated among other samples of (CLBP) pain patients. *Second*, given the restrictions in range on the TSK and RDQ, secondary to inclusion criteria, the construct validity correlations may actually underestimate the true relationships with PHODA-SeV, since range restriction attenuates correlations. The *third* limitation is that this study only used self-report measures, which may be influenced by self-presentational strategies. It would be worthwhile to test the construct validity of the PHODA-SeV by relating it to behavioural task performance. Because fear of movement/(re)injury, as measured by the TSK or Fear Avoidance Beliefs Questionnaire (Waddell et al., 1993), has consistently been found to be related to diminished task performance (Al-Obaidi et al., 2000; Geisser et al., 2000; Pfingsten et al., 2001; Al-Obaidi et al., 2003; Goubert et al., 2005b), replication of this with the PHODA-SeV would provide extended support for its construct validity.

Despite these limitations, it may be concluded that the reliability and validity of the PHODA-SeV are good to excellent. The decile scores (see Table 5) of the PHODA-SeV can provide researchers and clinicians with a frame of reference for the ratings of other patients with CLBP in rehabilitation and pain clinic settings. Although this is the first study providing such a frame of reference for PHODA-SeV scores, these data should be considered as preliminary, since the sample size on which the data are based is small. Further research is needed to establish normative data of the PHODA-SeV based on larger sample sizes. Another note of importance in interpreting these normative data is that these are based on a population with quite severe complaints. The participants suffered from CLBP for more than 10 years on average (see Table 1), were referred for rehabilitation care, and more than half were on sick leave or received disability pension. Yet, since the distribution of PHODA-SeV scores showed a large variation that covered almost the entire possible scale range, and since the duration of complaints did not influence the total score, it is possible that these normative score may also apply to chronic pain populations with less severe complaints.

This study used a short electronic version of the PHODA, and it can therefore be questioned whether these results can be generalized to the original PHODA. Because of the well-considered selection of 40 activities from the original ones, and because of the extremely high internal consistency, it is likely that these results extend to the complete PHODA. Moreover, due to the good to excellent properties of the PHODA-SeV, it can be suggested that this shortened edition may actually be more useful than the complete version.

In this study, the PHODA-SeV was administered as a measurement tool. It is thus unknown whether the results of this study generalize to the PHODA-SeV used for clinical purposes, during which ample interaction between the clinician and patient is possible. However, since the instruction and other basic properties of the PHODA were maintained during administration of the PHODA-SeV, there is no obvious reason to expect different results in applications as clinical diagnostic instrument.

## **ACKNOWLEDGEMENTS**

This study was supported by Grant nr. 1436.0002 of the Netherlands Organisation for Health Research and Development (ZonMw). We like to express our gratitude to Rosanne Jansen for programming the PHODA-SeV, and to Jeroen de Jong for selecting the items of the PHODA-SeV. Furthermore, we are grateful to Sita van Riet and Magda van Duinen for their assistance in the data collection, and to the therapists delivering the cognitive-behavioural treatments. We are also thankful to the two anonymous reviewers for providing valuable suggestions for improvement to the manuscript.

## **Chapter 6**

### **Measuring fear of movement/(re)injury in chronic low back pain using implicit measures**

**Maaïke Leeuw, Madelon L. Peters, Reinout W. Wiers, Johan W.S. Vlaeyen**

**This chapter is published as:**

Leeuw M, Peters ML, Wiers RW & Vlaeyen JWS. Measuring fear of movement/(re)injury in chronic low back pain using implicit measures. *Cognitive Behaviour Therapy* 2007; 36(1):52-64.



**ABSTRACT**

Fear of movement/(re)injury is assumed to contribute to the development and maintenance of chronic low back pain (CLBP) in a subgroup of patients. Studying fear of movement/(re)injury with implicit attitude measures, without the awareness of the patient, might be a valuable addition to self-report questionnaires. The aims of the current study were to investigate whether CLBP patients demonstrate more implicit fear of movement/(re)injury than healthy controls, and whether two implicit measures are related to each other, and to an explicit self-report measure of fear of movement/(re)injury. A group of 66 CLBP patients and 30 healthy controls took part in this study. In addition to self-report questionnaires, fear of movement/(re)injury was implicitly assessed by the Extrinsic Affective Simon Task (EAST) and the Go-No-Go-Association Task (GNAT) that aimed to determine the association between back-stressing movements and the evaluation 'threatening'. On both implicit tasks it was found that neither CLBP patients nor healthy controls demonstrated implicit fear of movement/(re)injury, and that CLBP patients did not differ from healthy controls in their level of implicit fear of movement/(re)injury. In general, no associations were found between the EAST and the GNAT, or between implicitly measured and self-reported fear of movement/(re)injury. One major caveat in drawing inferences from these findings is the poor reliability of these implicit measures. Research towards the psychometric properties of these measures should first be expanded before modifying, and applying, them to more complex domains such as fear of movement/(re)injury.

## INTRODUCTION

Of the many people suffering from low back pain (LBP), only a few will develop chronic low back pain (CLBP) complaints (Waddell, 2004c). A cognitive behavioural account assigns a prominent role to pain-related fear in explaining the development and maintenance of chronic low back pain in a subgroup of these patients (Lethem et al., 1983; Vlaeyen et al., 1995b). The detrimental contribution of pain-related fear in the development and maintenance of LBP complaints has been demonstrated in currently pain-free people (Linton et al., 2000; Picavet et al., 2002; Van Nieuwenhuyse et al., 2006), in acute LBP patients (Picavet et al., 2002; Burton et al., 2004; Swinkels-Meewisse et al., 2006c) and in CLBP patients (Geisser et al., 2000; Al-Obaidi et al., 2003; Goubert et al., 2005b). The effectiveness of treatments targeted at decreasing pain-related fear may also emphasize its contribution to disability in some patients (Vlaeyen et al., 2002b; Boersma et al., 2004; de Jong et al., 2005b). Pain-related fear is a rather general construct, encompassing various specific fears depending on the source to which the fear is directed. One of these specific fears, to which extensive research has been directed, is fear of movement/(re)injury that signifies the fear that physical activities and movements will cause (re)injury to the back.

Most research on pain specific beliefs is based on self-report measures, which may be susceptible to self-presentational strategies, and may be only suitable to measure beliefs that require a certain amount of conscious reflection and introspection (de Houwer, 2002; de Houwer, 2003b; Fazio and Olson, 2003). Conscious reflection might be lacking in fearful CLBP patients who often perceive their disability as a consequence of their pain rather than fear of movement/(re)injury (Vlaeyen et al., 2002c). Implicit attitude measures are suggested to be a promising method to overcome certain limitations of self-report measures, although it should be recognized that they should only be conceived as valuable additional tools and not as a replacement (e.g. de Jong et al., 2001; de Houwer, 2002; Gheldof et al., 2004). Implicit measures aim to measure evaluations or associations that are automatically activated without the awareness or control of the participant. They often consist of reaction time paradigms during which a response is required to words that are related to the association of interest. Implicit associations are assumed to be reflected in respectively facilitated or hampered task performance on trials that are compatible or incompatible to the underlying association (e.g. de Houwer, 2003b). Most studies report no, or only moderate, relationships between implicit and self-report measures (Fazio et al., 1995; Greenwald et al., 1998; Nosek and Banaji, 2001; Fazio and Olson, 2003; Teige et al., 2004). This discordance can be explained by the dual process MODE model (in which MODE is an acronym for *m*otivation and *o*pputunity as *d*eterminants of whether the attitude-to-behaviour process is primarily spontaneous or deliberative in nature; Fazio and Olson, 2003), through its proposition that responses to self-report measures are influenced by deliberative processes due to the presence of abundant motivation, time and recourses, while the responses of implicit measures rely on automatically activated processes because motivation and time are absent. Meta-analytic findings indicate that correlations between self-report and implicit measures systematically increased when either spontaneity in self-report measurement was

enhanced or when the conceptual overlap between these measures was high (Hofmann et al., 2005).

To date, only two studies have investigated implicit attitudes of CLBP patients, demonstrating ambiguous results. De Jong and Peters (2002) showed that highly fearful CLBP patients displayed stronger associations between pain and threat than did low fearful subjects on a Word Simon Task. Another study, using the Affective Priming Task, demonstrated a negative evaluation of back-stressing movements in pain-free participants, whereas no such negative evaluation was found for CLBP patients (Goubert et al., 2003). The investigated associations in these studies, however, do not reflect the main association we think underlies fear of movement/(re)injury, namely the evaluation of back-stressing movements as being threatening because of their potential to induce (re)injury to the back.

The current study aimed to examine implicit fear of movement/(re)injury in CLBP patients using two implicit measures that determined the association between back-stressing movements and the evaluation 'threatening'. The first task concerned the Extrinsic Affective Simon Task (EAST: de Houwer, 2003a), which was chosen because of its ability to measure different associations within the same task. Due to its relatively few applications, and low reliability (de Houwer, 2003a; Teige et al., 2004), the Go-No-Go Association Task (GNAT: Nosek and Banaji, 2001) was added because of its comparatively large effect size. First, we sought to investigate whether CLBP patients demonstrate more implicit fear of movement/(re)injury than healthy controls on the EAST and GNAT. Second, we determined whether these two implicit measures are related to each other, and to self-reported fear of movement/(re)injury.

## METHODS

### Participants

A sample of 66 CLBP patients and 30 healthy controls was included. Exclusion criteria were illiteracy, dyslexia, colour blindness, and a native language other than Dutch. The CLBP patients were selected from several sources (N = 38 from a randomized clinical trial of Maastricht University; N = 12 from a work reintegration service; N = 13 from participation in a previous study; N = 3 from responses to poster pamphlet). For all patients the main complaint concerned CLBP for more than three months (mean duration = 10.75 years, SD = 9.18, range between .5 and 38 years). Furthermore, they had to report at least some disability due to back pain, as indicated by a score of 3 or higher on the Roland Disability Questionnaire (Roland and Morris, 1983). The healthy controls were selected because of their participation in previous experiments of Maastricht University. They were included when they did not have a current pain problem or treatment, and had not experienced LBP complaints in the past year. This group was matched in age, gender, and educational level to the patient group.

### Procedure

The study was introduced as investigating differences in information processing between pain patients and pain free people. Verbal as well as written information about the study was provided. Informed consent was obtained before participation. All participants first performed the EAST and thereafter the GNAT. Subsequently, they

had to assess to what degree they considered the words of the implicit tasks as threatening. Finally, they completed several self-report questionnaires. Patients scheduled for treatment were tested prior to start of treatment, or in the first two weeks of treatment. All participants received a gift voucher and reimbursement of travel expenses for their participation.

## Materials

### Implicit measures

#### *EAST*

The EAST (de Houwer, 2003a) measures the strength of associations between certain concepts and attributes by means of a response time paradigm. Participants have to give a response ‘positive’ or ‘negative’ based on the colour of words, while the valence of that specific word has to be ignored. The underlying idea is that the performance should be superior on trials in which the valence of the correct response matches the meaning of the coloured word than on trials in which the valence of the word and the correct response are opposite to each other. For example, it should be easier to press the ‘positive’ key than the ‘negative’ key in response to ‘flower’, because of the universal positive evaluation of flowers (de Houwer, 2003a).

For the current study, the EAST was adapted to measure fear of movement/(re)injury. Instead of positive and negative evaluations that were used in the original EAST, and in the study of Goubert et al. (2003), the evaluations ‘threatening’ and ‘neutral’ were chosen as the most suitable associations with back-stressing movements. For exploratory reasons, rest words and injury words were also included in the task. Participants were requested to respond as fast and accurately as possible, by pressing the appropriate response key, to words that appeared in the middle of the computer screen. In the first practice block, white words (threatening and neutral) were presented, for which the correct response had to be based on the meaning of the word. Through this, one threatening and one neutral key were extrinsically generated. During the second practice block, coloured words (back-stressing movement words, rest words, and injury words) appeared in a green or blue colour, for which the response had to be based on the colour of the word. Coloured words therefore had to be responded to by either the threatening or neutral key, depending on the colour, and not the meaning, of the word. In the final actual test phase, these white and coloured words were presented interchangeably, with white words requesting a response based on their meaning and coloured words on the basis of their colour.

Implicit fear of movement/(re)injury should be reflected in a stronger association between back-stressing movements and the evaluation ‘threatening’ than ‘neutral’. From this time forth, associations corresponding to fear of movement/(re)injury will be referred to as *compatible* (pressing the threatening key in response to movement words), whereas associations non-equivalent to fear of movement/(re)injury will be referred to as *incompatible* (pressing the neutral key in response to movement words). By presenting each coloured word once in green and once in blue, for each word a compatible trial and an incompatible trial existed. An implicit association is reflected in better performance (faster response times and less errors) on compatible trials, and

in decreased performance (longer response times and more errors) on incompatible trials.

The first practice block consisted of 24 trials; six neutral (mineral, normal, objective, modern, owner, and adult) and six threatening (fatal, warning, dangerous, terrible, mean, and horrible) words were presented twice. The second practice block consisted of 36 trials; six back-stressing movements (bending, lifting, jumping, falling, pushing, running), six injury words (operation, disabled, damage, injury, stab, paralyzed) and six rest words (resting, sleeping, lying, relaxed, calmness, easiness) were presented once in green and once in blue. The test phase was composed of 60 trials; both the coloured and white words were presented twice, each coloured word once in green and once in blue. The words were presented after the presentation of a fixation dot during 500 ms. The response window was 2500 ms, after which the absence of a response was registered as a non-response. The inter trial interval was 1000 ms. When a response was shorter than 300 ms the message 'too fast' appeared on the screen. In case of an error, 'false' emerged in the screen until the accurate response was made. The assignment of the response keys to both the meaning and the colour of the words was counterbalanced between participants, resulting in four different EAST versions. For each version, the same word sequence was presented during the two practice blocks, while different word sequences were presented during the test phase of each version.

### *GNAT*

The GNAT (Nosek and Banaji, 2001) is another response time paradigm to measure implicit associations, by assessing the strength of an association between a certain concept and two attributes. During the GNAT several phases are presented. In each phase, different concepts are combined with different attributes. Due to this, both highly related and less related concepts and attributes are presented in different phases (for example, 'flower' can be combined with 'good' in one phase and with 'bad' in another). Participants have to press a key (GO) to stimuli belonging to pre-defined concept-attribute pairings, and do not have to press a key (NO GO) to stimuli that do not belong to either of these categories (for example, pressing the key to words belonging to the categories 'flower' and 'good', and not to 'bugs' and 'bad'). Performance should be superior when the key has to be pressed to a concept and attribute that are highly associated than when the key has to be pressed to a concept and attribute that are not, or less, associated with each other. For example, superior performance is expected when the key has to be pressed in case of items belonging to the categories 'flower' and 'good' than when the key has to be pressed to the categories 'flower' and 'bad'.

The GNAT was adapted to measure fear of movement/(re)injury. Back-stressing movement words and geometrical shape words represented the concept categories, while threatening words and neutral words represented the attribute categories. The shape phases were not relevant, but were included to present a neutral context. Participants had to decide as fast and accurately as possible whether they had to respond to words that appeared in the middle of the screen, by pushing a response key, or not to respond at all. The GNAT started with two practice tasks, in which participants had to differentiate between neutral words and threatening words, and between back-stressing movement words and geometrical shape words respectively.

After this, four actual test phases were presented, during which words from all four categories were presented intermixed. These test phases consisted of compatible and incompatible movement categorizations, as well as compatible and incompatible shape categorizations. The test phases of interest for this study concerned the compatible and incompatible movement phases that respectively indicated pressing the key to movement and threatening words, or to movement and neutral words. Based on the hits and false alarms, a sensitivity score was calculated according to the signal detection theory (Nosek & Banaji, 2001). The magnitude of the sensitivity index reflects the extent to which a concept is associated with an evaluation. Implicit fear of movement should thus be reflected in higher sensitivity on the compatible movement phase than on the incompatible movement phase.

During the first categorization phase 24 trials were presented: six neutral (mineral, normal, objective, modern, owner, adult) and six threatening (fatal, warning, dangerous, terrible, mean, horrible) words were presented twice. The second categorization phase also consisted of 24 trials: six movement (bending, lifting, jumping, falling, pushing, running) and geometrical shape words (square, triangle, rectangle, circle, oval, cube) were presented twice. In each test phase, the four categories of six words each were presented, resulting in 48 trials per test phase. The labels of the target categories were displayed in the right and left quadrant of the screen. The colour of both the stimuli and the category labels of the evaluative attributes was white, while the items and category labels of the concepts were displayed in light blue. The words were presented in the middle of the screen and remained there until the correct response was given and the feedback displayed, or until the response deadline was reached. The response window was 950 ms. Correct and incorrect responses were followed respectively by a green circle or a red cross that were presented for 500 ms. The next word appeared 250 ms after this feedback. Four versions of the GNAT were developed, in which the order of the compatible and incompatible phases was varied. For each phase of the GNAT, a different sequence of words was presented.

Both implicit tasks were programmed using micro experimental computer software (MEL; Schneider, 1996). Most of the presented words were derived from other studies (de Jong and Peters, 2002; Houben et al., 2005a; Vancleef, unpublished report). Others were devised while accounting for the concept of fear of movement/(re)injury as well as the length and number of syllables. All word sequences were generated by a randomization procedure, with the restriction that the same word was not presented directly again after its first presentation, and that the same response was not required for more than three times in a row. Responses were registered by an external response box (Psychology Software Tools, 1996-2003).

### Self-report measures

#### *Word appraisal*

Of every word that was presented during the implicit tasks, participants were asked to evaluate the threat value on a 10-point self-report rating scale ranging from 1 ('neutral') to 10 ('threatening'). Mean appraisal was computed for each of the word categories. Word appraisal was missing for one patient.

### *Functional Disability*

Functional disability was measured in all participants with the Dutch version (Schoppink et al., 1996) of the Quebec Back Pain Disability Scale (QBPDS; Kopec et al., 1995). The QBPDS consists of 20 items describing activities commonly affected by back pain, of which patients have to rate on a six-point self-report scale, ranging from 'not difficult' to 'unable to perform', to what degree they experience difficulties in performing these activities. The reliability and validity of the QBPDS are well established (Kopec et al., 1995; Schoppink et al., 1996; Fritz and Irrgang, 2001).

### *Fear of movement / (re) injury*

The CLBP patients completed the Dutch version (Vlaeyen et al., 1995c) of the Tampa Scale for Kinesiophobia (TSK; Miller et al., 1991) as a measure of fear of movement/(re)injury. The TSK consists of 17 statements that have to be rated on a 4-point scale ranging from 'strongly disagree' to 'strongly agree'. The Dutch version of the TSK has been shown to be sufficiently reliable and valid (Vlaeyen et al., 1995b; Swinkels-Meewisse et al., 2003a; Goubert et al., 2004d; Roelofs et al., 2004a). The healthy controls completed the TSK-G, in which all items have been rephrased in such way that people without low back pain can also complete them. The TSK-G was demonstrated to have sufficient internal stability and validity both for individuals with and without back pain (Houben et al., 2005b). For the TSK and TSK-G, a summation score was calculated after removal of the reversed key items because of too low inter-item correlations (Goubert et al., 2004; Houben et al., 2005).

### *Pain catastrophizing*

The Dutch version (Crombez and Vlaeyen, 1996) of the Pain Catastrophizing Scale (PCS; Sullivan et al., 1995) was used to determine catastrophizing about pain. The PCS consists of 13 items, on which the patients are asked to indicate on a five-point scale, ranging from 0 (not at all) to 4 (always), to what degree they experienced certain thoughts or feelings during pain. The Dutch version of the PCS shows good psychometric properties (Van Damme et al., 2000; Van Damme et al., 2002a).

## **Statistical analyses**

For each analysis, the skewness statistics of the included variables were examined. In cases where a variable was not normally distributed (skewness statistics below -1.5, or above 1.5), non-parametric analyses were employed.

The two practice blocks of the EAST and the white words of the actual test phase were removed from further analyses. EAST analyses were performed separately for the response latencies and the errors. Only response latencies of correct responses were included. In accordance with Greenwald et al. (1998) response latencies shorter than 300 ms. were recalculated into 300 ms. Log transformations were then performed on the raw latencies. Further analyses were performed with these transformed latencies, except in presenting mean response times. The proportion of errors was calculated by dividing the number of errors by the number of trials on which a response was given (thereby excluding non-responses). Both the mean response times and the mean proportion of errors were determined for the compatible and incompatible movement trials. Subsequently, an EAST response time score and an

EAST error score were calculated by subtracting the mean response latency or proportion of errors in the compatible phase from the mean response latency or proportion of errors in the incompatible phase. By this means, a positive EAST response time score or EAST error score indicated the presence of implicit fear of movement/(re)injury. Since associations with the movement words in the EAST were our main interest, the analyses focused on this word category. Data were separately analyzed for the response latencies and errors using analysis of variance (ANOVA) with a 2 (group: CLBP patients vs. healthy controls) x 2 (type of response: compatible movement trials or incompatible movement trials) design. Age, gender, educational level, and self-reported fear of movement/(re)injury were included as covariates.

For analyses with the GNAT, the first two practice blocks and the shape phases were removed from further analyses. Sensitivity indices were calculated for both the compatible and incompatible movement phase by the same method as described in Nosek and Banaji (2001). The GNAT movement effect was calculated by subtracting the sensitivity index of the incompatible movement phase from the one in the compatible movement phase. Due to this, a positive score would reflect the presence of implicit fear of movement. The same ANOVA was applied as with the EAST, except that for 'type of response' now 'phase' (compatible movement or incompatible movement) was included.

To investigate the relationship between implicit measures and self-report measures, analyses were separately performed for CLBP patients and healthy controls, since they were expected to differ. Pearson correlation coefficients were calculated to examine whether the EAST response time effect was related to the EAST error effect, and whether both EAST effects were related to the GNAT effect. Second, it was investigated whether implicit fear of movement/(re)injury measured by the EAST and the GNAT were related to self-report measures of fear of movement/(re)injury, by computing Pearson correlation coefficients with self-reported fear of movement/(re)injury, pain catastrophizing, and the appraisal of movement words.

## RESULTS

### Sociodemographics

The group descriptives for both the CLBP patients and the healthy controls are displayed in Table 1. Gender distribution, age, and education level did not differ significantly between patients and healthy controls. Patients reported significantly more functional disability, more fear of movement/(re)injury, and more pain catastrophizing than healthy controls. The total TSK score of the patients (without exclusion of the reversed key items) was 38.45 (SD = 6.44). Compared with normative data about the TSK (Vlaeyen et al., 1995b), it appeared that 57.6% of the patients scored below, or on the same level, as the reported median of the TSK. This suggests that the level of fear of movement/(re)injury of the included group of CLBP patients is representative for the average CLBP population.

### Stimulus words

To support the validity of the word selection, it appeared that the threatening words (mean = 7.36, SD = 1.78) were appraised as more threatening than the neutral words (mean = 1.40, SD = .65;  $Z = -8.46$ ,  $p < .001$ ). CLBP patients evaluated the movement



words of the implicit tasks as significantly more threatening (mean difference = -2.85,  $t = -8.11$ ,  $p < .001$ ) than healthy controls. Patients and healthy controls did not differ in their evaluation of neutral ( $U = 890.00$ ,  $p = .45$ ) or threatening (mean difference = -.24,  $t = -.61$ ,  $p = .54$ ) words. It therefore seems that the neutral and threatening words were accurately chosen for the key-assignment, as well as that the movement words were more threatening to the CLBP patients than healthy controls.

**Table 1**  
A summary of sample statistics

	CLBP patients (N = 66)	Healthy controls (N = 30)	Group difference statistics
Age (y)	43.88 ± 9.85	42.63 ± 8.68	$t = -.60$ , $p = .55$
Gender (% male)	53.03	50.00	$t = .27$ , $p = .79$
Education (%)*			$t = .72$ , $p = .48$
Low	25.76	0.00	
Middle-high	69.70	100.00	
Mean duration of low back pain (y)	10.75 ± 9.18	No LBP in the past year	-
QBPDS	47.70 ± 15.52	8.83 ± 10.18	$U = 66$ , $p < .001$
TSK(-G) without reversed key items	29.50 ± 5.94	23.03 ± 4.42	$t = -5.33$ , $p < .001$
PCS	17.62 ± 9.65	11.60 ± 8.41	$t = -2.95$ , $p < .01$

Note: QBPDS = Quebec Back Disability Scale; TSK(-G) = Tampa Scale for Kinesiophobia, either the version for the patient (TSK) or healthy control (TSK-G); PCS = Pain Catastrophizing Scale

\* Education: low = lower vocational education, or below; middle-high = advanced elementary education, or higher

**Implicit fear of movement/(re)injury in CLBP patients and healthy controls**

*EAST*

Mean proportion of errors was .11 (SD = .12) for CLBP patients, and .07 (SD = .07) for healthy controls. Mean response time was 862 (SD = 256) and 734 (SD = 220) respectively for CLBP patients and healthy controls. In accordance with De Houwer (2003a), three CLBP patients (error rates of 47% and 50%, and mean response time of 1741) and one healthy control (mean response time of 1649) were excluded from further analyses since either their response times or their error rates deviated more than three standard deviations from the mean of their reference group.

ANOVA with response latencies of the movement trials demonstrated no effect of type of response ( $F(1,86) = 1.22$ ,  $p = .27$ ), indicating that there was no difference in response times when back-stressing movements had to be responded to with either the threatening or the neutral key. Also, no interaction effect ( $F(1,86) = 1.64$ ,  $p = .20$ ) was found between group and type of response, implying that CLBP patients and healthy controls did not differ in their EAST response time effect (for CLBP patients mean = -.23, SD = 237.56); for healthy controls mean = 40.26, SD = 110.26). However, there was a significant effect of group ( $F(1,86) = 6.75$ ,  $p < .05$ ). Post hoc comparisons revealed that CLBP patients were significantly slower than healthy controls on the compatible ( $t = -3.28$ ,  $p < .01$ ) and incompatible ( $t = -2.49$ ,  $p < .05$ ) movement trials. In addition, the effect of age was found to be significant ( $F(1,86) = 7.46$ ,  $p < .01$ ). There was a positive correlation ( $r = .39$ ,  $p < .01$ ) between age and mean response time

on the movement words, indicating that older participants responded more slowly to these words than did younger participants.

Analyses with error rates on the movement trials also demonstrated no effect of type of response ( $F(1,86) = .00, p = .98$ ) or of the interaction between group and type of response ( $F(1,86) = .51, p = .48$ ). In addition, no group main effect was found ( $F(1,86) = .36, p = .55$ ). This indicates that there is neither a difference between errors on compatible and incompatible movement trials, nor between groups in EAST error effect (for CLBP patients mean =  $-.03, SD = .18$ ; for healthy controls mean =  $-.06, SD = .16$ ).<sup>2</sup>

### GNAT

In accordance with Nosek and Banaji (2001), five CLBP patients with sensitivity scores below zero on the movement phases were removed from further analyses. In addition, GNAT data of one patient were lost because of a computer problem.

ANOVA revealed no effect of phase ( $F(1,84) = .24, p = .63$ ), and no interaction effect of phase and group ( $F(1,84) = 1.05, p = .31$ ). This implies that there was no difference in sensitivity when the key was pressed to movement and threatening words as compared to movement and neutral words, neither overall, nor between groups. Therefore, it can be concluded that both groups did not differ with regard to their GNAT-effects (for CLBP patients mean =  $.71, SD = 1.17$ ; for healthy controls mean =  $.21, SD = 1.35$ ). However, a significant main effect of group was found ( $F(1, 84) = 5.41, p < .05$ ). Post hoc t-tests revealed that this was because patients had a significantly lower sensitivity on the incompatible phase ( $t = 2.84, p < .01$ ). Furthermore, significant effects of age ( $F(1, 84) = 6.96, p < .05$ ) and educational level ( $F(1, 84) = 4.34, p < .05$ ) were found. There were significant correlations between both age and educational level with sensitivity scores on the compatible ( $r = -.23, p < .05$  for age,  $r = .25, p < .05$  for educational level) and incompatible ( $r = -.31, p < .01$  for age,  $r = .26, p < .05$  for educational level) movement phases. A higher age and a lower educational level were therefore associated with diminished performance on the GNAT.

### Relationship between implicit measures and self-report measures

See Table 2 for an overview of the relationships between the implicit measures and self-report measures. Results showed that the EAST response time effect was not significantly related to the EAST error effect. Also, no significant relationship existed between both EAST effects and the GNAT effect. Thus, no significant relationship

<sup>2</sup> The same ANOVA applied to the response latencies and errors for the injury words revealed neither an effect of type of response ( $F(1,85) = 3.72, p = .06$  and  $F(1,86) = .37, p = .54$  respectively) nor an interaction between group and type of response ( $F(1,85) = 2.23, p = .14$  and  $F(1,86) = 2.57, p = .11$  respectively). While for the error rates none of the between subject variables was significant (all  $p > .12$ ), for the response times a significant effect was found for group ( $F(1,85) = 4.84, p < .05$ ) and age ( $F(1,85) = 11.23, p < .01$ ). Patients were significantly slower than healthy controls on the compatible injury trials ( $t = -2.50, p < .05$ ), and there was a positive correlation between age and mean response time on injury words ( $r = .36, p < .01$ ).

The same ANOVA applied to the response latencies and errors for the rest words revealed neither an effect of type of response ( $F(1,86) = .01, p = .92$  and  $F(1,86) = .37, p = .54$  respectively) nor an interaction between group and type of response ( $F(1,86) = .38, p = .54$  and  $F(1,86) = .06, p = .81$  respectively). While for the error rates none of the between subject variables was significant (all  $p > .10$ ), for the response times a significant effect was found for group ( $F(1,86) = 6.83, p < .05$ ) and age ( $F(1,86) = 7.79, p < .01$ ). Patients were significantly slower than healthy controls on both the compatible ( $t = -4.23, p < .001$ ) and incompatible ( $t = -2.45, p < .05$ ) rest trials, and there was a positive correlation between age and mean response time on rest words ( $r = .33, p < .01$ ).

was found between two different parameters within the EAST, nor between the EAST and the GNAT.

**Table 2**  
Correlation matrix of implicit measures and self-report measures for fear of movement/(re)injury for CLBP patients and healthy controls separately

		A	B	C	D	E	F
<b>Implicit measures</b>							
<b>A</b>	EAST effect <i>response time</i> , CLBP patients (N)	-					
	Healthy controls (N)	-					
<b>B</b>	EAST effect <i>error</i> , CLBP patients (N)	.05 (63)	-				
	Healthy controls (N)	-.08 (29)	-				
<b>C</b>	GNAT effect <i>sensitivity</i> , CLBP patients (N)	.20 (57)	.17 (57)	-			
	Healthy controls (N)	-.26 (29)	.05 (29)	-			
<b>Self-report measures</b>							
<b>D</b>	Fear of movement, CLBP patients (N)	-.01 (63)	.03 (63)	.07 (60)	-		
	Healthy controls (N)	.40 (29)*	-.21 (29)	.18 (30)	-		
<b>E</b>	Pain catastrophizing, CLBP patients (N)	.07 (63)	-.09 (63)	.11 (60)	.52 (66)***	-	
	Healthy controls (N)	.21 (29)	-.06 (29)	.16 (30)	.49 (30)**	-	
<b>F</b>	Word appraisal movement, CLBP patients (N)	.19 (62)	.27 (62)*	.09 (59)	.35 (66)**	.18 (66)	-
	Healthy controls (N)	-.08 (29)	.14 (29)	-.14 (30)	-.10 (30)	-.06 (30)	-

\* =  $p < .05$ ; \*\* =  $p < .01$ ; \*\*\* =  $p < .001$

Examination of associations between the implicit and self-report measures solely revealed a significant positive correlation between the EAST response time effect and self-reported fear of movement/(re)injury in healthy controls. Thus, the higher healthy controls scored on self-reported fear of movement/(re)injury, the higher they also scored on implicit fear of movement/(re)injury. The other significant association between the EAST error effect and appraisal of movement words for patients was considered irrelevant due to its weakness ( $r = .27$ ,  $p = .03$ ).

**DISCUSSION**

The main aim of the current study was to examine whether CLBP patients displayed more implicit fear of movement/(re)injury than healthy controls, as reflected in a stronger association between back-stressing movements and the evaluation ‘threatening’ being measured with the EAST and the GNAT. From analyses with both implicit measures, it appeared that neither CLBP patients nor healthy controls demonstrated implicit fear of movement/(re)injury. Furthermore, despite a significant difference on a self-report measure of fear of movement/(re)injury, CLBP patients did not differ from healthy controls in implicit fear of movement/(re)injury. Post-hoc examination of the injury and rest words, which were included in the EAST for exploratory reasons, corroborated these findings. Unfortunately, these results add even

more ambiguity to previous studies investigating implicit fear of movement/(re)injury (de Jong and Peters, 2002; Goubert et al., 2003).

The second aim of the study was to investigate whether the EAST and the GNAT were related to each other, as well as to self-report measures of fear of movement/(re)injury. The two different EAST parameters were not related to each other, nor did these two EAST parameters correlate with the GNAT score. These results are in accordance with other studies that failed to find any correspondence between implicit measures, which might be due to low to moderate reliability of implicit measures through which actual relationships between implicit measurements may be diminished (Bosson et al., 2000; Cunningham et al., 2001; Nosek and Banaji, 2001; Teige et al., 2004). It is also possible that differences in the designs of the two tasks are responsible for the lack of a relationship (Nosek and Banaji, 2001; de Houwer, 2003b). Although both measures aim to tap into the implicit association between movements and threat, they do this in rather different ways. *First*, the GNAT assumes faster responding when both movement and threat words should be responded to, whereas the EAST assumes faster responding to movement words with a key that has been extrinsically associated with threatening. *Second*, the GNAT compares performance between two different phases, while for the EAST comparison of performance within different trials is possible. *Finally*, different comparison categories were employed (e.g. physical shape in the GNAT, injury and rest words in the EAST).

In general, the implicit measures were also not related to the self-report measures. These findings are in line with the MODE model, since it can be expected that the applied implicit measures did not leave enough opportunity to consciously influence the response, thereby being discordant with responses to self-report measures that are consciously controlled (Fazio et al., 1995; Greenwald et al., 1998; Nosek and Banaji, 2001; Teige et al., 2004). The only potentially relevant relationship that appeared, namely between implicit and self-reported fear of movement/(re)injury in healthy controls, might also be explained by this model. The relative ease of healthy controls in performing these automatic association tasks may have left sufficient recourses to deliberately generate their responses to the EAST, resulting in higher overlap between the responses on the implicit and self-report measures.

Drawing inferences from these findings is rather intricate because of one major caveat of implicit measurements; that is, their questionable reliability and validity. Recently it became clear that reliability of the EAST is poor (de Houwer, 2003a; Teige et al., 2004). Indeed, we found that Cronbach's Alpha over the difference in response times between the compatible and incompatible presentation of each specific movement word, was extremely low ( $\alpha = .12$ ,  $N = 46$ ). For the GNAT, no information concerning reliability is available from existing literature. We examined the relationship between GNAT effects computed over the first and second half of the trials, which revealed a highly significant correlation ( $r = .50$ ,  $p < .001$ ;  $N = 89$ ) between both halves. This indicates moderate reliability. Due to the lack of adequate psychometric properties, especially for the EAST, it is possible that existing effects were not detected. Furthermore, for the current study we adapted the EAST and GNAT that were only previously applied to measure simple universal associations. This may also have reduced the reliability of these implicit measures.

There are several other limitations of this study that need to be taken into account. *First*, since CLBP patients considered these implicit tasks as more difficult to

perform (as reflected in their higher error rates, low sensitivity scores, and longer EAST response times) than healthy controls, it is possible that both groups experienced, or performed, these tasks differently. Through this, the implicit tasks might not be comparable between both groups. *Second*, participants performed the EAST prior to the GNAT. It is therefore possible, but unfortunately not verifiable, that the EAST influenced the performance on the GNAT. *Third*, medication use and pain intensity levels were not accounted for, though it is possible that these may have influenced the response times of the CLBP patients.

In conclusion, neither CLBP patients nor healthy controls demonstrated implicit fear of movement/(re)injury on the EAST and the GNAT. Neither implicit measure was related to the other, or to self-report measures of fear of movement/(re)injury. However, the poor psychometric properties of the implicit measures used to determine fear of movement/(re)injury should be acknowledged. Research encouraging the reliability and validity of these tasks may be warranted before they are modified and applied to complex domains, such as fear of movement/(re)injury.

## ACKNOWLEDGEMENTS

This study was supported by grant nr. 1436.0002 of the Netherlands Organisation for Health Research and Development (ZonMw). We are grateful to Sita van Riet of Maastricht University, and Floske de Jong, who realized the data collection. We also want to thank Dirk Hermans and two anonymous reviewers for their valuable comments on an earlier version of the manuscript. This study was made possible thanks to Jan Jochijms and Jolanda Idink of Top-Care Health Services Maastricht and Venlo.

## **Part IV**

### **Clinical implications of pain-related fear for chronic low back pain management**



## **Chapter 7**

### **Exposure in vivo versus operant graded activity in chronic low back pain patients: Results of a randomized controlled trial**

**Maaïke Leeuw, Mariëlle E.J.B. Goossens, Gerard J.P. van Breukelen,  
Jeroen R. de Jong, Peter H.T.G. Heuts, Rob J.E.M. Smeets,  
Albère J.A. Köke, Johan W.S. Vlaeyen**

**A subsequent version of this chapter is in press:**

Leeuw M, Goossens MEJB, van Breukelen GJP, de Jong JR, Heuts PHTG, Smeets RJEM, Köke AJA, Vlaeyen JWS. Exposure in vivo versus operant graded activity in chronic low back pain patients: Results of a randomized controlled trial. *Pain*, in press.



**ABSTRACT**

Since pain-related fear may contribute to the development and maintenance of chronic low back pain (CLBP), an exposure in vivo treatment (EXP) was developed for CLBP patients. We examined the effectiveness as well as specific mediating mechanisms of EXP versus operant graded activity (GA) directly and six months post-treatment in a multi-centre randomized controlled trial. In total, 85 patients suffering from disabling non-specific CLBP reporting at least moderate pain-related fear were randomly allocated to EXP or GA. It was demonstrated that EXP, despite excelling in diminishing pain catastrophizing and perceived harmfulness of activities, was equally effective as GA in improving functional disability and main complaints, although the group difference almost reached statistical significance favouring EXP. Both treatment conditions did not differ in pain intensity and daily activity levels either. Nor was EXP superior to GA in the subgroup of highly fearful patients. Irrespective of treatment, approximately half the patients reported clinically relevant improvements in main complaints and functional disability, although for the latter outcome the group difference was almost significant favouring EXP. Furthermore, the effect of EXP relative to GA on functional disability and main complaints was mediated by decreases in catastrophizing and perceived harmfulness of activities. In sum, this study demonstrates that up to six months after treatment EXP is an effective treatment, but not more effective than GA, in moderately to highly fearful CLBP patients, although its superiority in altering pain catastrophizing and perceived harmfulness of activities was clearly established. Possible explanations for these findings will be discussed.

## INTRODUCTION

Although low back pain (LBP) is a frequently encountered problem, only a small subgroup of patients with chronic complaints (CLBP) accounts for the huge economic consequences by the direct (e.g. medical consumption) and indirect (e.g. productivity loss due to sick leave) costs associated with pain and disability (van Tulder et al., 1995; Goossens, 2002). Since biomedical pathology cannot adequately explain complaints in most of these patients (Waddell, 2004b), cognitive-behavioural models propose that in some patients it may be pain-related fear that contributes to the development and maintenance of CLBP (Lethem et al., 1983; Philips, 1987; Vlaeyen and Linton, 2000). These models assume that when a person responds to pain with catastrophic interpretations about the cause and consequences of the pain, pain-related fear is likely to develop. The associated safety behaviours, such as avoidance and hypervigilance, may paradoxically maintain fear levels and worsen the pain and disability in the long run. Pain-related fear has been associated with the initiation of acute LBP (Picavet et al., 2002), and with the development and maintenance of CLBP (Swinkels-Meewisse et al., 2006c; Turner et al., 2006; for a review see Leeuw et al., 2007a).

CLBP patients usually receive physical or cognitive-behavioural treatments promoting function irrespective of their degree of pain-related fear (Ostelo et al., 2005; Vlaeyen and Morley, 2005). Fearful patients may benefit less from these therapies, since these may leave enough room for escape and avoidance behaviours (Al-Obaidi et al., 2005). For example, during graded activity (GA), which aims at restoring functional activities by positively reinforcing healthy behaviours, patients can avoid to select activities that they fear and therefore consider to be out of reach. Fearful patients might gain more profit from treatments that specifically target at decreasing pain-related fear. Recently, because of analogies with anxiety disorders (Vlaeyen and Linton, 2000; Leeuw et al., 2007a), an exposure in vivo treatment (EXP) has been developed especially for CLBP patients experiencing hindrance from their pain-related fear. By gradually exposing these fearful patients to previously avoided activities, their catastrophic misinterpretations can be corrected, leading to diminished pain-related fear and decreased disability. A number of studies using replicated single case experimental designs demonstrated that EXP was more effective than GA (Vlaeyen et al., 2001; Vlaeyen et al., 2002a; Vlaeyen et al., 2002b; de Jong et al., 2005b) or a no treatment baseline period (Linton et al., 2002a; Boersma et al., 2004). Yet, the effectiveness of EXP has to be determined in a larger sample of patients. Besides establishing the effectiveness of treatments, their postulated mechanisms of change should be determined as well (Kazdin and Nock, 2003; Vlaeyen and Morley, 2005; Turner et al., 2007).

The aim of this study was to examine the effectiveness, as well as specific mediating mechanisms, of EXP versus GA in CLBP patients referred for outpatient multidisciplinary rehabilitation care with at least moderate fear of movement/(re)injury by the application of a randomized controlled trial (RCT), directly and six months after treatment.

## METHODS

### Aims and hypotheses

Five hypotheses are formulated. Our primary hypothesis is that EXP is more effective than GA in reducing functional disability as well as main complaints. With respect to the secondary outcomes, EXP is expected to be more successful than GA in reducing perceived harmfulness of activities and pain catastrophizing, and in decreasing pain severity and increasing daily activity levels. Third, we expect that EXP results in more patients reporting clinically relevant improvements in functional disability and main complaints than GA. Fourth, we anticipate that especially for patients with heightened pain-related fear, EXP is more effective than GA, whereas smaller differences between treatment conditions are expected for those with lower levels of fear. Therefore, an interaction between treatment effect and baseline level of perceived harmfulness of activities is expected. Finally, we hypothesize that the effect of EXP as compared to GA is mediated by decreases in pain catastrophizing and perceived harmfulness of activities, which are the presumed mechanisms of change in EXP as compared to GA.

### Participants

CLBP patients applied for participation either by referral by physicians from various outpatient facilities in the Netherlands (two rehabilitation centres, four hospital departments of rehabilitation, one hospital department of anaesthesiology, and two departments of occupational physics), or by their response to an advertisement in a local newspaper. The inclusion criteria were: back pain for at least three months that was not caused by serious spinal injury (Waddell, 2004b), age between 18 to 65 years, the presence of a sufficient level of disability (Roland Disability Questionnaire  $> 3$ ), and the presence of at least moderate fear of movement/(re)injury (Tampa Scale for Kinesiophobia  $> 33$ ). Exclusion criteria were: illiteracy, pregnancy, substance abuse interfering with treatment, involvement in any litigation concerning the patients ability to work or disability income, specific medical disorders or cardiovascular diseases preventing participation in physical exercise, and serious psychopathology. Psychopathology was first determined with the SCL-90 (Arrindell and Ettema, 1986): total scores higher than 241, corresponding to the upper 5% of the Dutch norms for chronic pain patients, were considered to signal interfering psychopathology. Furthermore, scores higher than 14 on the subscale 'negative self view' (Morley et al., 2002) of the Beck Depression Inventory (Beck et al., 1961) were considered indicative of depressive disorder. All patients were seen by a rehabilitation physician, who excluded patients with a diagnosis of serious spinal injury (Waddell, 2004b), and who checked whether rehabilitation care was suitable. A research assistant verified the remaining selection criteria. Patients were requested not to seek diagnostic or therapeutic procedures during therapy other than their allocated treatment, but were permitted to continue their prescribed medication. Patients who met all selection criteria were included in the study after providing their informed consent. To ensure equal expectations, patients were informed that both treatments were currently used in the participating treatment centres, and that they both aimed at restoring functional activity levels, but that their effectiveness had not yet been compared. The treatment names were not revealed (they were called 'BackActive-E' and 'BackActive-G'), to

ensure the study be single-blinded. At the start of their treatment, patients received information about their allocated treatment by their therapist(s). Only after completion of the long-term follow-up assessment of the study, they received additional information about the contrasting treatment. Treatments were carried out by mini-teams consisting of a psychologist, and a physiotherapist or occupational therapist, who were supervised by a rehabilitation physician in one of four outpatient rehabilitation centres in the Netherlands (Rehabilitation Centre Hoensbroeck, Rehabilitation Centre Blixembosch, department of rehabilitation of Maastricht University Hospital, department of rehabilitation of Maasland Hospital Sittard). In case patients were referred to the study from another location than where the treatments were carried out, they were treated at the centre of their preference.

The Medical Ethics Committee of Maastricht University/University Hospital Maastricht approved the research protocol in addition to the institutional committees of the participation institutions. The trial also received a trial identification number ([www.controlled-trials.com](http://www.controlled-trials.com); ISRCTN88087718).

### **Randomization procedure**

Patients were randomized to EXP or GA within each treatment centre, following a predetermined and computer-generated randomization schedule, prestratified by therapist team, and degree of pain catastrophizing and disability. Patients were assigned to the next available therapist team. For disability and pain catastrophizing, the median score of the data from a previous RCT was used as the cut-off (van den Hout et al., 2003). Within each stratum, a randomized block design with a block size of two was used. By this, it was first ensured that important patient characteristics were equally distributed across conditions, and second, that each therapist team would approximately treat an equal number of patients in each treatment condition. Because of the complexity of the randomization schedule, forthcoming assignments could never be predicted. The randomization schedule was only accessible for the research assistant performing the randomization. Treatment assignment was concealed for participants until they arrived at their first intervention session. After the second pre-treatment measurement, patients received a sealed envelop from the research assistant containing a sheet of coloured paper indicating treatment assignment, which they opened together with the psychologist during the intake.

### **Interventions and therapists**

Although EXP and GA share a number of commonalities, they also have important unique features.

#### *Common features of EXP and GA*

Both treatments are aimed at restoring functioning and decreasing limitations, though EXP aims to achieve this goal by systematically reducing pain-related fear by utilizing Pavlovian conditioning and cognitive therapeutic techniques (Vlaeyen et al., 2004a), whereas GA aims to increase healthy behaviour by employing operant learning principles (Fordyce, 1976; Fordyce, 2001b; Fordyce, 2001a). Both treatment conditions start with a psychological intake aimed to establish a cognitive and behavioural analysis of the pain problem. They also include an educational session by the rehabilitation physician prior to start of treatment. During this session, the results

of the medical examination are discussed, the findings of X-rays are explained when available, and explicit assurance is provided that being active, rather than taking rest, is beneficial to the back. Additionally, in both conditions, the therapists extensively explain the treatment rationale to the patients, and provide them with a written summary of this rationale. Finally, patients are encouraged to formulate feasible treatment goals in terms of specific activities previously confined, and treatment sessions are provided twice a week in one-hour sessions, with a decreasing frequency during the final weeks of treatment to increase generalization.

### *Unique features of EXP*

EXP aims at improving functional ability by reducing the perceived harmfulness of activities. The treatment is structuralized and consists of approximately 16 sessions in total. After the psychological intake (one session), the treatment starts with the establishment of a personal graded hierarchy of fear-eliciting activities by use of the Photograph Series of Daily Activities (PHODA; Kugler et al., 1999): patients have to organize 100 photographs of daily activities in ascending order based on the extent to which they believe that performing these activities would be harmful to the back (one session). After this, two educational sessions are provided, of which the first is carried out by the rehabilitation physician, and the second by the therapist mini-team explaining the treatment rationale by integrating the individual complaints and characteristics of the patient in the fear avoidance model. Subsequently, patients are gradually but systematically exposed to personally tailored and fear-provoking activities that are derived from the individual fear hierarchy of the PHODA. Besides, behavioural experiments are performed, during which patients can test the validity of their beliefs about the harmful consequences of activities, and can construct and test new, more adaptive beliefs (approximately twelve sessions). The therapist mini-team carried out these treatment sessions. A more detailed description of EXP is provided elsewhere (Vlaeyen et al., 2002c; Vlaeyen et al., 2004a).

### *Unique features of GA*

GA aims at improving functional ability by positively reinforcing health behaviours and activity levels. The treatment is also structuralized and consists of approximately 26 sessions in total. After the psychological intake procedure (one session), the treatment rationale is explained, emphasizing the detrimental effects of immobility and beneficial consequences of activity (two sessions, of which one by the rehabilitation physician, and the other by the therapist mini-team). Treatment goals are identified in terms of specific functional activities by means of the Patient Specific Complaints (Beurskens et al., 1996), from which patients can select their three to five most important functional activities confined by their pain problem. During the pain-contingent baseline assessment, the tolerance level of these activities is determined by instructing patients to engage in activities until their limit of tolerance (approximately six sessions). This is followed by the establishment of a time-contingent treatment schedule in which activities should be initially performed at 70% to 80% of the baseline activity level, and then gradually increased in time towards the preset treatment goals (approximately fifteen sessions). Patients are instructed not to deviate from these quotas, not even when they feel capable of doing so. In addition, positive reinforcers are provided when the patient accomplishes the predefined activity quotas.

Most practice takes place at home in between treatment sessions, whereas during the sessions the patient's progress towards the goals is evaluated. Patients are instructed to keep graphical registrations of their performance. Two additional sessions with a psychologist take place to help spouses with positively reinforcing activity levels, thereby enhancing the generalization process. Operant behavioural graded activity is shown to be superior to no treatment (Lindstrom et al., 1992; Morley et al., 1999; Ostelo et al., 2005; Smeets et al., 2007). Except for the psychological sessions, a physiotherapist or occupational therapist carried out the treatment sessions. The details of GA are described elsewhere (e.g. Sanders, 2002).

### *Therapists*

All participating therapists were affiliated with either of the four treatment centres, except for one psychologist who was appointed by the study, and who operated in three of the sites. This psychologist was involved in the treatment of 46 out of 85 patients (54%), whereas the other psychologists only contributed between 1 and 18 % of the treatments delivered. All therapists had at least half a year experience in treating CLBP patients in rehabilitation care. Different combinations of psychologists with physiotherapists or occupational therapists resulted in 19 mini-teams, which treated between one and twelve patients each. The teams always performed both EXP and GA, and were therefore not blinded to the contents of the treatments being compared. They were extensively trained in both treatments, and were provided with detailed treatment manuals that gave specific instructions for each session, and included guidelines for dealing with problems during the treatment process. Therapists attended to collective supervision sessions that took place three times a year.

### **Protocol adherence and contamination check**

Research comparing treatment outcomes should aim to assess the fidelity of the treatments delivered to ensure honest and genuine comparisons (Yeaton and Sechrest, 1981; Vermilyea et al., 1984; Moncher and Prinz, 1991; Borrelli et al., 2005). It was therefore verified whether both treatments were carried out according to the treatment protocols (sufficient protocol adherence), and whether these were not contaminated with elements from the contrasting treatment or other proscribed elements (absence of contamination). A random sample of 30 sessions out of 265 that were recorded by the therapists was drawn (11% of the available recordings), while taking care that for each therapist team sessions from both treatments were selected when possible. Any information explicitly revealing the treatment condition was removed from the recording.

A self-developed instrument measured protocol adherence and treatment contamination. This measure consisted of several specific treatment elements for which their occurrence or non-occurrence during treatment delivery had to be indicated without revealing whether these were required, allowed, or prohibited to occur. Additionally, it requested to judge to which treatment condition the session belonged. The specific treatment elements of this measure were identified by two experts of both treatment protocols (J.V. and M.L.). The content validity of this instrument was supported by sufficient independent agreement between these experts ( $\kappa = .73$ ), which was corroborated by the trained therapists (73%), in categorizing these elements into five categories (1. essential and unique; 2. essential but not unique;

3. unique but not essential; 4. compatible; and 5. prohibited; which were slightly adapted from Pereplechikova and Kazdin, 2005, and Waltz et al., 1993). Two undergraduate students in clinical psychology, who were kept blind to the study hypotheses, and who were trained and paid for that purpose, evaluated the recordings of the treatment sessions using this instrument.

### **Treatment credibility/expectancy**

Treatment outcome can be influenced by the degree to which patients consider a treatment as credible, and to what extent patients expect favourable outcome for themselves (Kole-Snijders et al., 1999; Goossens et al., 2005; Smeets et al., *In press*). It was therefore determined whether the credibility and expectancies were equal for both treatment conditions, using a Dutch translation and modification of the credibility-expectancy questionnaire (CEQ; Devilly and Borkovec, 2000), that will further be denoted as CEQ-m (the Dutch version of the CEQ-m is presented in Appendix A). Two factors (credibility and expectancy) of the original CEQ have shown high internal consistency and adequate test-retest reliability (Devilly and Borkovec, 2000; Smeets et al., *In press*). After the explanation of the treatment rationale, participants completed the CEQ-m, assessing credibility and expectancy on the domains of disability as well as daily activity levels and pain severity (11 items), which could be answered on a nine-point scale ranging from ‘totally not’ to ‘totally’ (total scores ranging between 11 and 99).

In addition to patients’ credibility of treatment, the degree to which therapists believe in the treatment they have to deliver may also be of importance (Wampold, 2000). After their training in both treatments, therapists had to rate how credible they considered both EXP and GA as treatments for CLBP patients with pain-related fear, with eight items of the credibility scale of the CEQ-m (total scores ranging between 8 and 72; the Dutch version of this credibility scale of the CEQ-m for therapists is presented in Appendix B).

### **Measures**

Assessments occurred twice before treatment (Pre-treatment 1 and 2), and were repeated directly after termination of treatment (Post-treatment), and six months (Follow-up 1), and twelve months (Follow-up 2) afterwards. The current paper only reports on the results until six months follow-up. In accordance with the intention-to-treat principle, not only patients who completed the study, but also those who withdrew from treatment were invited to perform measurements after their withdrawal. All assessments were performed at the test locations (Maastricht University and Rehabilitation Centre Blixembosch). The two research assistants collecting the data were not blinded to treatment allocation, but all questionnaires were administered electronically through a computer program. Patients not willing to travel to the test centre were requested to complete the assessment by paper and pencil at home, for whom consequently only self-report measures were available. Patients received reimbursement of travel expenses for the assessments they attended. In addition to the measures presented below, various other questionnaires were administered, as well as daily assessments during treatment and a cost diary to measure LBP associated costs. Since these measures are irrelevant within the scope of this study, these will not be discussed here.

### *Baseline assessment*

Before treatment, data was gathered about gender, age, educational level, employment status, duration of complaints, previous treatments including low back surgery, and pain radiation to the leg. Additionally, fear of movement/(re)injury was assessed with the Tampa Scale for Kinesiophobia (TSK: Miller et al., 1991; Vlaeyen et al., 1995a; Roelofs et al., 2007), and depressive symptoms with the Beck Depression Inventory (BDI: Beck et al., 1988).

### *Primary outcomes*

To assess the severity of *functional disability*, the Dutch version (Schoppink et al., 1996) of the Quebec Back Pain Disability Scale (QBPDS; Kopec et al., 1995; Kopec et al., 1996) was administered. Patients are requested to rate their difficulties with performing 20 listed activities commonly affected by back pain on a six-point rating scale ranging from 'no trouble' to 'unable to' (total score ranging between 0-100). The QBPDS has a high internal consistency, good test-retest reliability (Kopec et al., 1995; Schoppink et al., 1996; Davidson and Keating, 2002; Mousavi et al., 2006), sufficient responsiveness (Davidson and Keating, 2002) and validity (Reneman et al., 2002). We prefer the QBPDS to the commonly used Roland Disability Questionnaire (RDQ: Roland and Morris, 1983), because the QBPDS is especially developed for LBP patients from a conceptual approach to disability measurement, whereas the RDQ is not (Kopec et al., 1995; Schoppink et al., 1996). Other advantages are that its content most closely resembles our outcome of interest, and that it may be more sensitive and responsive (Kopec et al., 1995; Schoppink et al., 1996; Gommans et al., 1997).

Second, *main complaints* were assessed by the Patient Specific Complaints (PSC; Beurskens et al., 1996). The PSC lists various physical activities, out of which patients can select their three main individual complaints. Patients rate their difficulties with performing these activities during the past week on a 100-millimeter visual analogue scale, ranging from 0 ('no problem at all') to 100 ('impossible'). The same selection of activities was assessed on all measurement occasions.

### *Secondary outcomes*

First, *perceived harmfulness of activities* was determined with the short electronic version of the Photograph Series of Daily Activities (PHODA-SeV: Leeuw et al., 2007b). The PHODA-SeV consists of 40 photographs of daily activities that are successively presented on a computer screen. Patients have to imagine performing the particular activity displayed on the picture, and to indicate to what degree they consider this activity to be harmful to the back by placing the photograph along a 'harmfulness thermometer' ranging from 0 ('not harmful at all') to 100 ('very harmful'). A mean total score ranging from 0 to 100 is calculated as the sum of each rating divided by 40. The PHODA-SeV has been shown to be internally consistent, to have good test-retest reliability, and to be valid (Leeuw et al., 2007b).

Second, *pain catastrophizing* was assessed with the Pain Catastrophizing Scale (PCS: Sullivan et al., 1995), which has good psychometric properties (Osman et al., 2000; Van Damme et al., 2002a). It consists of 13 items, on which the patients are asked to indicate on a 5-point scale, ranging from 0 (not at all) to 4 (always), to which degree they experience certain thoughts or feelings during pain.



Third, *daily activity level* was objectively measured in the natural environment of the patients. Patients were requested to wear the RT3 Tri-axial Research Tracker (Stayhealthy), which is a lightweight battery-powered ambulatory activity monitor that can be worn clipped to the waistband during waking hours for approximately eight days. A composite score of acceleration signals was calculated over the sampled and accumulated counts on three individual axes (vertical, anteroposterior, and mediolateral) over one-minute epochs. The accelerometer was not worn during activities involving water (e.g. swimming, taking a shower), or in case of possible damage to the device (e.g. basketball). The RT3 is a valid and reliable measure of activity levels (DeVoe et al., 2003; Powell et al., 2003; Rowlands et al., 2004).

Finally, although both treatments were not aimed at pain reduction, behavioural therapies have been shown to affect pain intensity levels (Morley et al., 1999; Ostelo et al., 2005) as well as some single case studies comparing EXP with GA (Vlaeyen et al., 2001; Vlaeyen et al., 2002b). Therefore, the final secondary outcome was *pain intensity*, which was measured with three 100 mm visual analogue scales ranging from 'no pain at all' to 'unbearable pain' of the McGill Pain Questionnaire (MPQ; Melzack, 1975) measuring current pain intensity, the most intense pain, and the least intense pain. A pain intensity score was calculated as the mean of these three ratings.

### *Clinically relevant improvements*

We used two measures to assess clinically relevant improvements. The first one, the PSC measuring main complaints, is also one of the primary outcomes, for which the responsiveness has been documented by Beurskens et al. (1996). Second, since for the RDQ data are available to examine clinical improvements (Roland and Morris, 1983; Beurskens et al., 1996; Gommans et al., 1997), whereas for the QBPDS there are not, we added this questionnaire to the outcome measures (Stratford et al., 1998; Roland and Fairbank, 2000). The RDQ consists of 24 questions about the occurrence of functional limitations due to LBP that can be answered by either yes or no. Its test-retest reliability (Roland and Morris, 1983), responsiveness to change (Grotle et al., 2004a; Rocchi et al., 2005), and validity (Reneman et al., 2002; Rocchi et al., 2005) are supported.

### **Sample size**

Although the previous single case studies establishing the effectiveness of EXP (Vlaeyen et al., 2001; de Jong et al., 2002; Linton et al., 2002a; Vlaeyen et al., 2002b; Boersma et al., 2004) employed the RDQ whereas we utilized the QBPDS as the main outcome measure, the same effect sizes were presumed for both outcomes. These single case studies reported a mean post-treatment difference of 7.6 points on the RDQ. Since these studies did not find any effect of GA, derived from data regarding improvements in GA patients from an RCT of van den Hout (2002) with similar participants, we expected a mean difference between conditions of 4.37 points (within-group SD = 6.46, implying an effect size  $d = 0.68$ ). Given Alpha = 5% (two-sided) and power = 90%, statistical significance would be reached with 47 patients in each condition. Accommodation for possible dropout at follow-up (15%) resulted in a requirement of 55 patients per condition.

### Statistical analyses

Results are reported according to the Consolidated Standards of Reporting Trials (CONSORT: Moher et al., 2005; Plint et al., 2006). Statistical analyses were performed with SPSS statistical software, version 13.0 (SPSS, Inc., Chicago).

#### *Treatment groups at baseline*

The mean of both pre-treatment measurements was calculated for each measure and patient, and used in subsequent analyses, since none of the variables showed a significant change between the two pre-treatment assessments (p-values ranging between .15 and .93 based on paired t-tests). To verify whether randomization resulted in equal treatment groups, we tested pre-treatment scores between the treatment conditions by means of independent t-tests and chi-square tests ( $\alpha = .10$ ).

#### *Protocol adherence and contamination*

In case of sufficient interrater reliability of the instrument to measure protocol adherence and contamination (Cohen's kappa  $> .70$ ), the mean scores of the evaluators were used in subsequent analyses. In the absence of previous guidelines, the cut-off points for sufficient protocol adherence and absence of treatment contamination were judiciously, but arbitrarily, determined. Protocol adherence was first regarded as sufficient when the percentage of accurate treatment classifications exceeded 90%, and second, when at least 70% of the required treatment elements ('essential and unique', and 'essential but not unique') had actually occurred during the treatments. Contamination was considered to be absent when no more than 10% of prohibited treatment elements ('proscribed') took place. We verified by ANOVA whether the mean protocol adherence and contamination scores computed over all rated treatment recordings were equal for both treatment conditions.

#### *Treatment credibility/expectancy*

For the patients ( $N = 72$ ), the factor structure of the CEQ-m was examined with principal axis factor analysis with Oblimin Rotation. For each factor found, Cronbach's Alpha was computed to assess its internal consistency. We verified whether the treatment groups were equal in their mean credibility-expectancy with independent t-tests in case of adequate psychometric properties. For the therapists ( $N = 27$ ), Cronbach's Alpha was computed to assess the internal consistency of the credibility items, separately for both treatment conditions. It was furthermore tested whether the therapists considered GA equally credible as EXP using a paired t-test.

#### *Treatment outcome: intention-to-treat analyses*

Intention-to-treat analyses were employed, including all patients that were originally enrolled in the study, irrespective of them completing therapy or assessments (Hollis and Campbell, 1999). We performed mixed (multilevel) regression, which uses all available data without requiring imputation of missing data (Little, 1995; Verbeke and Molenberghs, 2000; Moerbeek et al., 2003). The dependent variables were the primary and secondary outcome measures. Our basic design was a repeated measures design with group (EXP, GA) as a between-subject factor and time of measurement as within-subject factor. When baseline variables differed significantly ( $p < .10$ ) between treatment conditions despite randomization, these were controlled for by including

them into the outcome analyses as covariates. Adjustments were furthermore made for age, gender, educational level, and treatment centre. Also some potentially relevant prognostic variables measured at baseline were entered into the model, which were duration of complaints, financial compensation status, and perceived harmfulness of activities<sup>3</sup>. We confirmed normality and homogeneity of variance of the residuals as well as the absence of outliers or influential cases for each outcome. Each test of significance consisted of a likelihood ratio (LR) test with  $df = k$ , with  $k$  being the difference in parameters between the two models being compared. Before starting analyses with the initial mixed model, we verified for each covariate separately whether it interacted with treatment. None of these interactions were significant, as were none of the group by covariate interaction for any covariate at any time point ( $\alpha = .01$  or  $\alpha = .005$  for the primary and secondary outcomes respectively).

In Appendix C, the initial mixed model that we used is explained with an equation, which may be helpful to understand the modelling steps and final results. This initial model consisted of all main effects and group by time and covariate by time effects, which was then reduced to improve parsimony and interpretability by the following steps. *First*, non-significant covariate by time interactions (*covariate\*post* and *covariate\*foll* in Appendix C) were removed from the model using  $\alpha = .01$  or  $\alpha = .005$  for the primary and secondary outcomes respectively (these restrictive  $\alpha$ 's were used in view of multiple testing). *Second*, all main effects of covariates not involved in a covariate by time interaction were stepwise deleted if not significant using the same  $\alpha$ 's. *Third*, the main group effect ( $\beta_1$  in Appendix C) was dropped, since this is the difference between treatments at baseline, which must be zero due to randomization, and this increases the power of the group by time effect of interest (Laird and Wang, 1990; van Breukelen, 2006). *Fourth*, for every significant covariate by time interaction, equality of covariate effects at both time points was tested for all covariates simultaneously ( $\alpha = .05$ ). If equality was confirmed, the model was simplified accordingly. *Finally*, equality of the group difference at both time points was tested ( $\alpha = .05$ ). By these steps, only statistically significant prognostic variables remained in the model, and were therefore controlled for in the estimate of the treatment effect. For the treatment effect,  $\alpha$  was set to .05 for the primary outcomes and .01 for the secondary ones.

For the secondary outcome physical activity level, some preparatory statistical processing was necessary with Matlab 6.5.1. The first day of the accelerometry data was always excluded because it was incomplete, leaving seven potential days of recording. The time interval between the start and end of activity reflected the duration of data recording per day. A 'standard measurement day' was determined, which was the length of time that at least 70% of the participants had worn the monitor. Then, 'complete days' were identified, which were those with activity registration over at least 80% of a standard measurement day (Catellier et al., 2005). Only patients with no less than five complete days, including at least one weekend day, were included in the analyses (Gretebeck and Montoye, 1992). A mean daily activity level was calculated

<sup>3</sup> Categorical covariates were entered using dummy coding, with gender: 0 for females, 1 for males; financial compensation: 1 for presence of and 0 for absence of financial compensation; centre1: 1 for Maastricht, 0 for others; centre2: 1 for Hoensbroek, 0 for others; centre3: 1 for Sittard, 0 for others; education-low: 1 for low education, 0 for middle and high education; education-high: 1 for high education, 0 for middle and low education. Furthermore, each quantitative covariate was standardized to prevent collinearity due to the inclusion of interaction terms.

by dividing the total amount of activity registered by the number of days of registration. Since it appeared from data examination that one accelerometer was deviant, which was from an older type than the other devices, data from this device were excluded from further analyses.

#### *Treatment outcome: per-protocol analyses*

In addition to intention-to-treat analyses, per-protocol analyses were performed, incorporating only those patients that finished their treatment, to compare effectiveness of EXP versus GA on the premise that a patient completes treatment. The same analyses as according to the intention-to-treat principle were employed with respect to the primary outcomes.

#### *Clinically relevant improvements*

Change scores based on all available data were computed between pre- and post-treatment and pre-treatment and follow-up 1 for the PSC and the RDQ. For the PSC, a decrease of 21 mm. on the mean of the three main complaints was considered a clinically relevant improvement (Beurskens et al., 1996). For the RDQ, patients with pre-treatment scores between 0-4, 5-8, 9-12, 13-16, 17-20, 21-24 had to improve at least 2, 2, 4, 5, 8, and 8 points respectively (Stratford et al., 1998; Roland and Fairbank, 2000). By means of logistic regression analyses on both outcomes separately for post-treatment and follow-up assessment, including group and all covariates of the mixed regression models (see 2.9.4) as independent variables, it was verified whether the proportion of clinically improved patients differed between treatment conditions.

#### *Interaction of treatment with perceived harmfulness of activities*

Given that we included both moderately and highly fearful patients, of interest was testing the interaction between perceived harmfulness of activities and treatment, to check whether the treatment effect was different for highly fearful patients than for patients with a low fear level ( $\alpha = .01$  or  $\alpha = .005$  for the interaction test with respect to primary and secondary outcomes respectively). When equality of this interaction effect at both time points was confirmed ( $\alpha = .05$ ), the interaction was presented accordingly.

#### *Specific mediating mechanisms*

Mixed regression according to the intention-to-treat principle was employed. Mediation was examined with regard to the primary outcomes by the joint significance test, which was put forward by MacKinnon (2002) as a straightforward method that outperformed many other mediation tests in terms of power and type I errors (Pituch et al., 2005; Mallinckrodt et al., 2006). First, the effect of treatment on the mediators was already examined with mixed regression following the procedure outlined in section 2.9.4, since the mediators 'pain catastrophizing' and 'perceived harmfulness of activities' were also secondary outcomes. Second, we tested the effect of the mediator on the primary outcomes QBPDS and PSC, adjusted for treatment, by adding the mediator as time-dependent or within-subject covariate to the final mixed model for the QBPDS and PSC. So the baseline mediator value served as covariate for the baseline outcome measurement, the post-test mediator value as covariate for the post-test outcome, and the follow-up 1 mediator value as covariate for the follow-up 1

outcome measure. The joint significance of the association between treatment and mediator, and the association between the mediator and outcome adjusted for treatment ( $\alpha = .05$ ), provides evidence for mediation, assuming that the mediator affects the outcome and not vice versa.

## RESULTS

### Flow of participants

Figure 1 shows the flow of participants including withdrawals and reasons at each stage of measurement or treatment. The study took place between September 2003, when the first participant was presented to the study, and January 2007, when the final follow-up 1 measurement was completed (the final follow-up 2 assessment is planned to be completed in July 2007). In total, 177 CLBP patients applied for participation, of whom 138 were referred to the study by physicians from participating rehabilitation facilities and 39 responded to a newspaper advertisement. Since 20 patients already refrained from partaking in the screening, 157 patients were actually assessed for eligibility. Of these, 85 (54%) satisfied the selection criteria, and were included in the study. Randomization resulted in 42 patients being allocated to EXP and 43 to GA, who all completed both pre-treatment measurements.

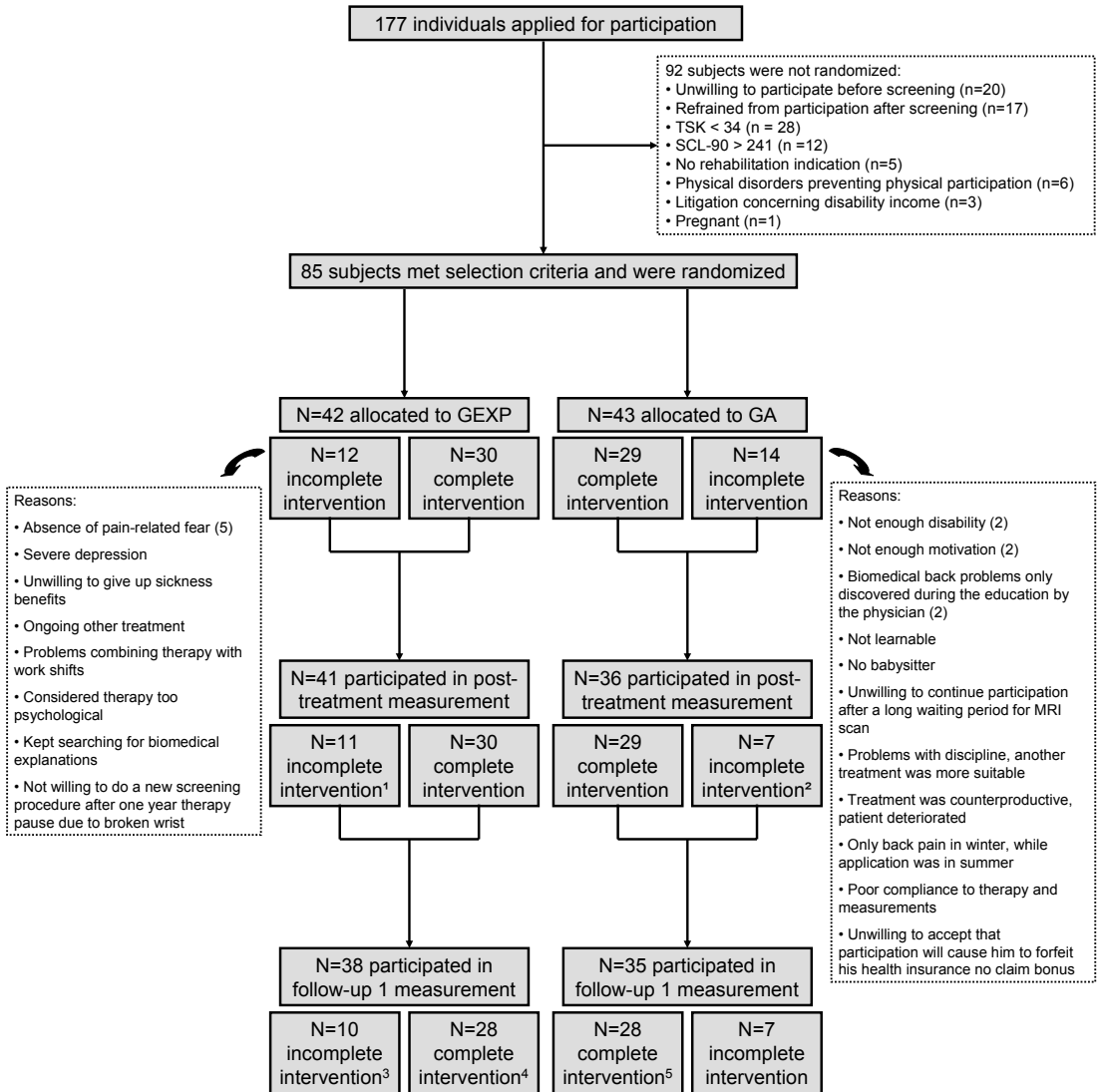
Treatment was prematurely terminated either by the patient or the therapist in twelve patients (29%) of the EXP condition, and in fourteen patients (33%) of the GA condition for various reasons described in Figure 1. As can be derived from these, none of the patients reported any adverse or side effects related to the interventions, except for one GA patient who deteriorated during treatment. EXP patients received 12.55 (SD = 5.16) treatment sessions on average (for completers  $M = 15.10$ , SD = 2.75 and for dropouts  $M = 6.17$ , SD = 4.15), and GA patients 17.53 (SD = 8.73) on average (for completers  $M = 22.72$ , SD = 3.02 and for dropouts  $M = 6.80$ , SD = 6.47 respectively).

Of those patients completing treatment, in four EXP patients and one GA patient treatment consisted of fewer sessions than prescribed, since their treatment was so successful that continuation no longer was deemed appropriate. There were also protocol deviations in three other completers because of insufficient understanding of the Dutch language (EXP), concurrent treatment (GA), and interruption due to severe illness in the family (GA). In total, eight patients (9%) refused to participate in the post-treatment assessment (EXP: 1, GA: 7), and twelve patients (14%) refused to complete follow-up 1 measurement (EXP: 4, GA: 8). As can be derived from Figure 1, the proportion of missing data is acceptable, and the reasons for non-response do not seem to be related to the content of the treatments (e.g. adverse effects).

### Treatment groups at baseline

Table 1 displays the baseline variables for the total sample and per intervention group. The groups did not differ on the primary outcomes (all  $p \geq .10$ ), the secondary outcomes (all  $p > .36$ ), the outcomes assessing clinically relevant changes (all  $p \geq .10$ ), or on the presumed mediators (all  $p > .36$ ). Both groups were also equal in demographic variables, back pain complaints, and work status (all  $p > .10$ ), although unequal baseline scores in depressive symptoms were found ( $p = .03$ ).

**Figure 1**  
Flow of participants



- <sup>1</sup> Refused to repeat screening and pretreatment measurements since previous data became invalid after too long interruption of treatment due to broken wrist
- <sup>2</sup> Lack of time; Rejected participation after long waiting period for additional diagnostic tests; Research team was notified only several months after termination of treatment; Rejected treatment and study (4).
- <sup>3</sup> Lack of time
- <sup>4</sup> Unreachable (2)
- <sup>5</sup> Rejected measurements

**Table 1**  
Summary of demographic characteristics, work status, complaints, baseline questionnaires, and the outcomes measures of the total sample of participants and the two treatment groups

	Total (N = 85)	EXP (N = 42)	GA (N = 43)	p <sup>1</sup>
Age (y)	45.32 ± 9.45	46.45 ± 9.33	44.21 ± 9.54	.28
Gender (% male)	51.8 %	47.6 %	55.8 %	.45
Education (%)				.42
Low	43.5 %	45.2 %	41.9 %	
Middle	44.7 %	47.6 %	41.9 %	
High	11.8 %	7.1 %	16.3 %	
Work status (%) <sup>2</sup>				
Work	35.3 %	31.0 %	39.5 %	
Sick leave	28.2 %	28.6 %	27.9 %	
Disability pension	25.9 %	33.3 %	18.6 %	
Other (e.g. study, household, unemployment)	30.6 %	31.0 %	30.2 %	
Duration of LBP (y)*	9.00 ± 9.40	8.13 ± 9.95	9.00 ± 8.94	.92
Pain radiation to the legs (%)	97.6 %	97.6 %	97.7 %	.99
Use of medication (%)	71.8 %	76.2 %	67.4 %	.37
Previous operation(s) on back	30.9 %	36.6 %	25.6 %	.38
Previous treatment(s) for LBP	98.8 %	100.0 %	97.7 %	.32
Functional disability, QBPDS (0-100)	52.74 ± 12.59	53.61 ± 11.63	51.88 ± 13.54	.53
Main complaints, PSC (0-100)*	75.67 ± 14.90	71.50 ± 14.94	77.67 ± 14.69	.10
Harmfulness of activities, PHODA-SeV (0-100)	53.24 ± 18.51	51.35 ± 19.84	55.08 ± 17.16	.36
Pain catastrophizing, PCS (0-52)	22.93 ± 10.44	23.05 ± 10.57	22.81 ± 10.44	.92
Daily activity level, RT3 <sup>3</sup>	218110 ± 87673 (N=60)	210859 ± 83196 (N=33)	226972 ± 93677 (N=27)	.48
Pain intensity, MPQ (0-100)	53.61 ± 12.30	52.54 ± 12.77	54.66 ± 11.88	.43
Functional disability, RDQ (0-24)	14.74 ± 3.55	15.23 ± 3.64	14.27 ± 3.44	.22
Fear of movement/(re)injury, TSK (4-68)	42.01 ± 6.15	42.18 ± 7.21	41.85 ± 4.99	.81
Depressive symptoms, BDI (0-63)*	11.00 ± 7.48	12.75 ± 7.46	9.50 ± 7.30	.03
Treatment credibility/expectancy, CEQ-m (11-99)* <sup>4</sup>	77.00 ± 12.93 (N=72)	74.00 ± 13.39 (N=35)	77.00 ± 12.59 (N=37)	.57

Values presented are means ± SD, or percentages, except when marked with \*, when medians are presented since skewness exceeded -1 or 1.

EXP = exposure in vivo; GA = graded activity; LBP = low back pain; QBPDS = Quebec Back Pain Disability Scale; PSC = Patient Specific Complaints; PHODA-SeV = Short electronic Version of the Photograph Series of Daily Activities; PCS = Pain Catastrophizing Scale; RT3 = Stayhealthy RT3 accelerometers; MPQ = McGill Pain Questionnaire, mean of three VAS measuring current pain, least pain intensity, and maximal pain intensity; RDQ = Roland Disability Questionnaire; TSK = Tampa Scale for Kinesiophobia; BDI = Beck Depression Inventory ; CEQ-m = modified version of the Credibility Expectancy Questionnaire

<sup>1</sup>  $\chi^2$  tests, independent t-tests or Mann-Whitney U tests for comparisons between treatment groups with  $\alpha = .10$ .

<sup>2</sup> The categories of work status were not mutually exclusive, so more than one category could apply for each participant. Due to this, chi-square test was inappropriate.

<sup>3</sup> For n=25 data were not included in the analyses: no accelerometers were available at the test location of Rehabilitation Centre Blixembosch (n=7), data were lost due to defective devices (n=2), the registered data did not meet the criteria for inclusion in the analyses (n=10), or the data were registered with the deviant accelerometer (n=6).

<sup>4</sup> For n=13 no CEQ-m data were available, either because the treatment was terminated before this questionnaire had to be completed (n=9), or because the therapists had forgotten to administer this questionnaire (n=4).

### Protocol adherence and contamination check

Interrater reliability between both observers was good, with Cohen's kappa = .87 for the assessment of treatment condition, and .72 for the specific treatment elements. Both evaluators classified the sessions correctly in terms of treatment condition in 97% of the cases: each evaluator incorrectly classified one EXP session as GA. This supports the differentiation of both treatments. After dichotomization of the contamination sum scores due to an extremely non-normal distribution, no differences were found in both protocol adherence or contamination scores between both treatment conditions ( $F = .11$ ,  $p = .74$  and  $\chi^2 = 3.07$ ,  $p = .08$  respectively). It was shown that on average 72% (SD = 18.71) of the essential elements occurred during actual delivery of both treatments ( $M = 73.17$ ,  $SD = 16.97$  for EXP, and  $M = 71.53$ ,  $SD = 20.55$  for GA), whereas on average only 4% (SD = 7.71) of the prohibited elements took place ( $M = 5.72$ ,  $SD = 9.24$  for EXP, and  $M = 2.36$ ,  $SD = 5.44$  for GA).

### Treatment credibility/expectancy

The CEQ-m items completed by the patients measured a single factor, which explained 61% of the variance, on which all items loaded significantly (ranging between .47 and .86). Its internal consistency was excellent (Cronbach's Alpha = .93). The use of the total score is therefore justified for further analyses. No differences were found in credibility/expectancy ratings between both treatment conditions (Mann-Whitney test,  $p = .57$ ).

The credibility items of the CEQ-m that were completed by the participating therapists showed good internal consistency (Cronbach's Alpha = .81 and .87 for EXP and GA respectively). Prior to start of the trial, the therapists considered EXP ( $M = 55.60$ ,  $SD = 6.34$ ) as a more credible treatment than GA ( $M = 45.08$ ,  $SD = 11.08$ ;  $t = 4.29$ ,  $p < .001$ ).

### Treatment outcome: intention-to-treat analyses

Table 2 displays the means of the primary and secondary outcomes for both treatment groups at each time point. Since mixed regression confirmed equality of the group difference (EXP-GA) on both time points for all primary and secondary outcome measures, the estimated group differences were combined for post-treatment and follow-up 1 assessment.

The estimated group differences are displayed in Table 3, as are their effect sizes. No significant difference was found between treatment conditions in functional disability ( $p = .09$ ) and main complaints ( $p = .08$ ) directly and six months after treatment, although the group difference almost reached significance favouring EXP. With regard to the secondary outcomes, it was found that the EXP group had significantly lower PHODA-SeV ( $p < .001$ ) and PCS ( $p < .01$ ) scores directly and six months after treatment than the GA group, indicating that the manipulation of perceived harmfulness and pain catastrophizing was successful in EXP. No group difference was found for daily activity levels<sup>4</sup> ( $p = .73$ ), or for pain intensity levels ( $p = .90$ ), directly or six months post-treatment.

<sup>4</sup> As can be derived from Table 2, activity data of several participants were unavailable or excluded from analyses. Logistic regression with inclusion/exclusion of daily activity data as dependent variable for the 3 assessments separately and the variables of the subsequent mixed regression as covariates showed that the patients whose data were included in the analyses did not differ from those whose data were excluded ( $p$ -values  $> .02$  for pre-test;  $p$ -values  $> .09$  for post-treatment;  $p$ -values  $> .23$  for follow-up 1). Also, available daily activity data was equal for both treatment conditions ( $\chi^2 < 2.55$ ,  $p > .11$  for all measurement occasions).



**Table 2**

Observed means and standard deviations (SD) based on all available data for the scores on the primary and secondary outcome measures by condition at pre-treatment, post-treatment, and six months follow-up 1 assessment

Primary outcomes	Pre-treatment Mean ± SD	Post-treatment Mean ± SD	Follow-up 1 Mean ± SD
Functional disability, QBPDS (0-100)			
EXP	53.61 ± 11.63 (N=42)	35.90 ± 20.45 (N=41)	39.00 ± 20.93 (N=38)
GA	51.88 ± 13.54 (N=43)	41.69 ± 22.58 (N=36)	41.94 ± 19.29 (N=35)
Main complaints, PSC (0-100)			
EXP	68.82 ± 14.94 (N=42)	40.24 ± 29.42 (N=41)	38.52 ± 26.71 (N=38)
GA	73.36 ± 14.69 (N=43)	45.98 ± 27.49 (N=36)	44.06 ± 28.48 (N=35)
Secondary outcomes	Pre-treatment Mean ± SD	Post-treatment Mean ± SD	Follow-up 1 Mean ± SD
Harmfulness of activities, PHODA-SeV (0-100)			
EXP	51.35 ± 19.84 (N=42)	20.97 ± 17.65 (N=40) <sup>1</sup>	26.35 ± 23.32 (N=32) <sup>2</sup>
GA	55.08 ± 17.16 (N=43)	42.07 ± 21.07 (N=35) <sup>1</sup>	41.63 ± 21.26 (N=31) <sup>2</sup>
Pain catastrophizing, PCS (0-52)			
EXP	23.05 ± 10.57 (N=42)	13.73 ± 10.24 (N=41)	14.66 ± 10.80 (N=38)
GA	22.81 ± 10.44 (N=43)	17.53 ± 12.02 (N=36)	15.34 ± 9.41 (N=35)
Daily activity level, RT3			
EXP	210859 ± 83196 (N=33) <sup>3</sup>	213680 ± 76326 (N=26) <sup>4</sup>	232636 ± 88402 (N=24) <sup>5</sup>
GA	226973 ± 93678 (N=27) <sup>3</sup>	247229 ± 94213(N=21) <sup>4</sup>	224057 ± 79093 (N=19) <sup>5</sup>
Pain intensity, MPQ (0-100)			
EXP	52.54 ± 12.77 (N=42)	43.72 ± 21.24 (N=41)	41.15 ± 22.26 (N=38)
GA	54.66 ± 11.88 (N=43)	44.07 ± 22.86 (N=36)	40.45 ± 22.25 (N=35)

EXP = exposure in vivo; GA = graded activity; QBPDS = Quebec Back Pain Disability Scale; PSC = Patient Specific Complaints; PHODA-SeV = Short electronic Version of the Photograph Series of Daily Activities; PCS = Pain Catastrophizing Scale; RT3 = Stayhealthy RT3 accelerometers; MPQ = McGill Pain Questionnaire, mean of three VAS's measuring current pain intensity, least pain intensity, and maximal pain intensity

<sup>1</sup> For n=1 from EXP and n=1 from GA no PHODA-SeV data were available because they completed only questionnaires of the post-treatment assessment at home

<sup>2</sup> For n=6 from EXP and n = 4 from GA no PHODA-SeV data were available because they completed only questionnaires of the follow-up 1 assessment at home

<sup>3</sup> For n=25 data were not included in the analyses: no accelerometers were available at the test location of Rehabilitation Centre Blixembosch (n=7), data were lost due to defective devices (n=2), the registered data did not meet the criteria for inclusion in the analyses (n=10), or the data were registered with the deviant accelerometer (n=6)

<sup>4</sup> For n=38 data were not included in the analyses: no accelerometers were available at the test location of Rehabilitation Centre Blixembosch (n=7), refusal to complete post-treatment assessment at all (n=8), only willing to complete questionnaires of post-treatment assessment at home (n=2), the registered data did not meet the criteria for inclusion in the analyses (n=17), or the data were registered with the deviant accelerometer (n=4)

<sup>5</sup> For n=42 data were not included in the analyses: no accelerometers were available at the test location of Rehabilitation Centre Blixembosch (n=7), refusal to complete post-treatment assessment at all (n=12), only willing to complete questionnaires of post-treatment assessment at home (n=3), the registered data did not meet the criteria for inclusion in the analyses (n=17), or the data were registered with the deviant accelerometer (n=3)

**Table 3**

Estimated group difference (B) and 95% confidence intervals (C.I.) in primary and secondary outcomes at post-treatment and follow-up 1 assessment based on intention-to-treat analyses

	B <sup>1</sup> (SE)	95% C.I.	p	ES <sup>2</sup>	95% C.I.
<b>Primary outcomes</b>					
Functional disability (QBPDS) <sup>a</sup>	-5.89 (3.41)	-12.68 to .90	.09	-.31	-.67 to .05
Main complaints (PSC) <sup>b</sup>	-10.03 (5.64)	-21.27 to 1.21	.08	-.39	-.82 to .05
<b>Secondary outcomes</b>					
Harmfulness of activities (PHODA-SeV) <sup>c</sup>	-17.98 (3.80)	-25.56 to -10.40	<.001	-.90	-1.29 to -.52
Pain catastrophizing (PCS) <sup>d</sup>	-5.20 (1.81)	-8.82 to -1.59	<.01	-.53	-.89 to -.16
Daily activity level (RT3) <sup>e</sup>	-.07 (.20)	-.46 to .33	.73	-.07	-.47 to .33
Pain intensity (MPQ) <sup>f</sup>	-.56 (4.35)	-9.24 to 8.11	.90	-.03	-.42 to .37

QBPDS = Quebec Back Pain Disability Scale; PSC = Patient Specific Complaints; PHODA-SeV = Short electronic Version of the Photograph Series of Daily Activities; PCS = Pain Catastrophizing Scale; RT3 = Stayhealthy RT3 accelerometers; MPQ = McGill Pain Questionnaire, mean of three VAS's measuring current pain intensity, least pain intensity, and maximal pain intensity

<sup>1</sup> Since EXP is coded as 1, and GA as 0, a negative B indicates that EXP demonstrated lower scores at post-treatment and follow-up 1 than GA. The B's reported are the B of the group by time interaction effect in the model in Appendix C, where time=1 for post-test and follow-up, and time=0 for baseline, and where group by time replaces the terms group by post and group by foll since these were found to have the same B-value (i.e. B5=B6).

<sup>2</sup> The effect size is Cohen's d, calculated by dividing B by the square root of the average of the residual variances at post-test and follow-up.

<sup>a</sup> Adjusted for main effects of gender, duration of complaints, financial compensation status, harmfulness of activities, depressive symptoms, post-test, follow-up 1, and the interaction of duration of complaints by time, and harmfulness of activities by time.

<sup>b</sup> Adjusted for main effects of duration of complaints, financial compensation status, harmfulness of activities, depressive symptoms, post-test, follow-up 1, and the interaction of duration of complaints by time, harmfulness of activities by time, and depressive symptoms by time.

<sup>c</sup> Adjusted for main effects of duration of complaints, post-test, follow-up 1.

<sup>d</sup> Adjusted for main effects of financial compensation status, harmfulness of activities, depressive symptoms, post-test, follow-up 1, and the interaction of financial compensation status by time, and harmfulness of activities by time.

<sup>e</sup> Adjusted for main effects of gender, post-test, follow-up 1.

<sup>f</sup> Adjusted for main effects of gender, duration of complaints, harmfulness of activities, post-test, follow-up 1, and the interaction of duration of complaints by time, and harmfulness of activities by time.

### Treatment outcome: per-protocol analyses

Analyses including only patients who completed their treatment corroborated the results found with intention-to-treat analyses for the QBPDS (B = -6.92, SE = 3.81, p = .08). For the PSC, the patients completing EXP treatment reported borderline significant lower pre-treatment PSC scores than the patients completing GA treatment (B = -9.78, SE = 4.14, p = .02), and no difference between treatment conditions was found directly and six months after treatment when corrected for this baseline group difference by including the group effect into the model of Appendix C (B = -1.72, SE = 7.58, p = .82).

### Clinically relevant improvements

The baseline PSC and RDQ scores, and subsequent changes from pre-treatment to post-treatment and follow-up 1 in both treatment conditions, are displayed in Table 4, as are the number and proportions of patients reporting clinically relevant improvements. No significant group difference was found between treatment conditions in the number of patients reporting clinically relevant changes in main complaints (B = -.05, p = .92 and B = .94, p = .10 for post-treatment and follow-up 1

respectively), or in functional disability ( $B = 1.00$ ,  $p = .08$  and  $B = .98$ ,  $p = .07$  for post-treatment and follow-up 1 respectively), although for the latter outcome the group difference almost reached significance favouring EXP.

**Table 4**  
Means and standard deviations (SD), and the proportion of patients reporting clinically relevant changes, based on all available data for the change in the outcomes for clinically relevant change from pre-treatment to post-treatment and follow-up 1

	Pre-treatment Mean ± SD	Mean change pre- treatment to post-treatment ± SD	Mean change pre-treatment to follow-up 1 ± SD
Main complaints, PSC (0-100)			
EXP	68.82 ± 14.94	-28.41 ± 31.42 (N=41)	-30.11 ± 30.92 (N=38)
GA	73.36 ± 14.69	-27.70 ± 29.24 (N=36)	-29.30 ± 28.07 (N=35)
Functional disability, RDQ (0-24)			
EXP	15.23 ± 3.64 (N=42)	-6.35 ± 5.60 (N=41)	-6.34 ± 5.75 (N = 38)
GA	14.27 ± 3.44 (N=43)	-4.40 ± 6.13 (N=36)	-4.23 ± 5.60 (N = 35)
<b>Proportion of patients reporting clinically relevant change</b>		<b>Post-treatment</b>	<b>Follow-up 1</b>
Main complaints, PSC (0-100) <sup>1</sup>			
EXP		21/41 (51%)	24/38 (63%)
GA		21/36 (58%)	18/35 (51%)
Functional disability, RDQ (0-24) <sup>2</sup>			
EXP		22/41 (54%)	19/38 (50%)
GA		15/36 (42%)	12/35 (34%)

EXP = exposure in vivo; GA = graded activity; PSC = Patient Specific Complaints; RDQ = Roland Disability Questionnaire

<sup>1</sup> A decrease of 21 mm. on the mean of the three VAS scales assessing main complaints was considered as a clinically relevant improvement.

<sup>2</sup> Patients with pre-treatment scores between 0-4, 4-8, 13-16, 17-20, 21-24 had to improve at least 2, 2, 4, 5, 8, and 8 points respectively to consider the change as a clinically relevant improvement.

**Interaction of treatment with perceived harmfulness of activities**

For none of the outcomes was a statistically significant interaction found between treatment and baseline level of perceived harmfulness of activities ( $p = .46$  for the QBPDS;  $p = .60$  for the PSC;  $p = .08$  for daily activity level;  $p = .78$  for pain intensity), indicating that the difference between treatment groups did not depend on the level of pain-related fear.

**Specific mediating mechanisms**

In section 3.5, it was already shown that treatment was significantly related to the presumed mediators, indicating that EXP was more effective than GA in reducing pain catastrophizing and perceived harmfulness of activities ( $p < .01$  and  $p < .001$  respectively). Inspection of the relation between the mediators and outcome when controlling for treatment revealed that both pain catastrophizing ( $B = .43$ ,  $SE = .09$ ,  $p < .001$  for the QBPDS;  $B = .70$ ,  $SE = .15$ ,  $p < .001$  for the PSC) and perceived harmfulness of activities ( $B = .29$ ,  $SE = .04$ ,  $p < .001$  for the QBPDS;  $B = .28$ ,  $SE = .07$ ,  $p < .001$  for the PSC) were significantly related to the primary outcomes. The effect of treatment on the primary outcomes after inclusion of the mediators was only

marginal ( $B = .15$ ,  $SE = 2.59$ ,  $p = .96$  for the QBPDS;  $B = -1.52$ ,  $SE = 4.61$ ,  $p = .74$  for the PSC), and much smaller than those without adjustment for the mediators (see Table 3). This conclusion about mediation is not at variance with the failure to find a significant treatment effect on either primary outcome before inclusion of the mediators (section 3.5), since McKinnon et al. (2002) showed that testing of this treatment effect may suffer from low power if mediation is complete (i.e. if there is almost no direct treatment effect on the outcomes in addition to the mediated effect). In conclusion, the effect of EXP relative to GA on functional disability and main complaints was significantly mediated by decreases in pain catastrophizing and perceived harmfulness of activities.

## DISCUSSION

This study failed to demonstrate that EXP is a more effective treatment than GA for CLBP patients reporting at least moderate fear of movement/(re)injury. Despite its superior ability in diminishing patients' perceived harmfulness of activities and pain catastrophizing, EXP was about equally effective as GA in improving functional disability and main complaints directly and six months after treatment, although the group difference almost reached significance favouring EXP. Both treatments were also equally effective in reducing pain intensity and increasing daily activity levels. Per protocol analyses did not demonstrate that EXP would be more effective than GA on the presumption that patients complete their treatment. Additionally, we found no indication that EXP is superior to GA in the subgroup of highly fearful patients, thus both treatments are of equal value irrespective of the patient's level of fear. Furthermore, the difference between EXP and GA with respect to functional disability and main complaints was mediated by decreases in pain catastrophizing and perceived harmfulness of activities. Although the current study design does not allow for causal inferences with respect to the relation between the mediators and the outcomes, these findings are in line with the assumption that the effectiveness of EXP in reducing disability and main complaints relative to GA might be due to diminishment in pain catastrophizing and perceived harmfulness of activities.

Irrespective of treatment condition, approximately half the patients reported clinically relevant improvements in main complaints and functional disability, although for the latter outcome the difference between groups almost reached significance favouring EXP. Lacking a waiting list control group, we compared our findings with a Dutch RCT with similar patients and a similar treatment context (Smeets et al., 2006b), that found GA with problem solving skills training to be more effective than a waiting list control group (observed mean pre-post treatment improvement of  $M = 2.65$ ,  $SD = 4.66$  on the RDQ). The finding that in our study larger changes were found for GA as well as EXP ( $M = 4.40$ ,  $SD = 6.13$  and  $M = 6.35$ ,  $SD = 5.60$  respectively) given comparable baseline levels, obviously establishes the effectiveness of both treatments in clinically improving functional disability reports at post-treatment.

Since each therapist team participated in both treatment conditions, we verified protocol adherence and treatment contamination in both treatments. According to our preset criteria applied to evaluations of recorded treatment sessions, sufficient differentiation and protocol adherence came forward as well as absence of contamination. Most importantly, the equal protocol adherence and contamination for

both treatments suggests that neither of the treatments was favoured over the other. Consequently, comparisons of treatment effectiveness seem fair. Some caution, however, is warranted. First, it is possible that, despite adequate protocol adherence, treatment competence is compromised (Waltz et al., 1993), as are the understanding and utilization of the treatment skills by the patients (Bellg et al., 2004). Furthermore, only 21% of all treatment sessions were recorded, so it may be that therapists only recorded the treatment sessions for which they expected high protocol adherence and low contamination, and abstained from recording in case of sessions for which they expected the opposite, due to which treatment fidelity assessment may be inflated (Moncher and Prinz, 1991; Perepletchikova and Kazdin, 2005).

So why did we fail to find that EXP is a more effective treatment than GA? *First*, this study may suffer from a lack of power, as we were not able to include the number of patients that were required (85 instead of 110), and the effect sizes found were smaller than anticipated in our sample size calculation. *Second*, it may be that GA comprised some valuable components that were also part of EXP. Foremost, GA may also have included some form of exposure in vivo, since patients performed various activities, including those that were fear eliciting. In addition, GA also included the explicit assurance that all activities were safe and allowed, as was communicated during the educational sessions of the rehabilitation physician and the therapist mini-team, which may have resulted in modifications of patients pain-related beliefs (de Jong et al., 2005b). Finally, because the therapists were also trained in EXP, their attitude towards activities may have changed accordingly in the sense that they were less careful and worried about possible activity restrictions for their patients (Linton et al., 2002b; Coudeyre et al., 2006). *Third*, GA may have contained several favourable aspects as compared to EXP, except for the therapists ascribing higher credibility to EXP than to GA. GA consisted of more treatment sessions (26) than EXP (16), and patients selected and performed activities that were of high importance during GA, whereas patients had to perform activities that were selected based on their threat value during EXP. Furthermore, since the therapists were already accustomed to carrying out GA, whereas the majority of them had no previous experience, nor gained much practice, with EXP, it may be that specific treatment competence was superior in GA. *Fourth*, meta-analyses on psychotherapy outcome studies showed that treatment outcome can be mainly explained by common treatment factors, such as therapeutic competence or the patient's and therapist's belief in the effectiveness of the treatment, while specific treatment ingredients only contribute marginally (Wampold, 2000). The fact that EXP and GA shared more common features than that they comprised unique treatment ingredients might explain the insignificant differences between the treatments.

The unadjusted effect sizes for the difference between treatment conditions, obtained by running the final mixed models minus all covariates, were  $-.24$  and  $-.18$  for functional disability and main complaints respectively. These correspond to the mean effect size reported by meta-analyses comparing (cognitive)-behavioural treatments with active control interventions (Morley et al., 1999; Hoffman et al., 2007). Notwithstanding, the results of the current study do not corroborate the results of the preceding single case studies showing the discernible superiority of EXP in comparison to GA, possibly because in these studies differentiation between treatment conditions was artificially enlarged by restricting GA to only low threatening

activities, and because of high specific competence of the therapist delivering EXP (Vlaeyen et al., 2001; Vlaeyen et al., 2002a; Vlaeyen et al., 2002b; de Jong et al., 2005b).

Approximately one out of three patients withdrew from both treatments. This number of treatment withdrawals is somewhat higher than others reported (Mannion et al., 1999; Smeets et al., 2006b), but comparable to Woby et al (2004b). For none of these patients was the reason to terminate treatment related to the specific content of that treatment, except for five EXP patients for whom the therapists decided that fear of movement/(re)injury was absent. From these experiences, it may be concluded that EXP may not be suitable in the absolute absence of fear of movement/(re)injury.

The study sample seems representative for the general rehabilitation population, so the results are likely to generalize to other patients being referred for rehabilitation in the Netherlands. Our sample reported somewhat more severe functional disability in comparison to CLBP patients referred for rehabilitation care (Reneman et al., 2002). Furthermore, the duration of complaints was high in comparison to the study of Smeets et al. (2006b), but equal to Kole-Snijders (1999). On the other hand, the percentage of patients receiving financial compensation was about the same as compared to these studies, and the levels of fear of movement/(re)injury were also in line with other reports (Goubert et al., 2000).

Some drawbacks of this study need mentioning. *First*, the statistical processing of the accelerometer data can be debated. We refrained from complex analyses regarding interruption in wearing time, or days on which patients were sick and therefore abstained from wearing the device, which were therefore not accounted for in the activity estimations. However, we have no reason to believe that the number and duration of interruptions in wearing or days of sickness differed between treatment groups, neither before treatment because of randomization, nor after treatment because both treatments aimed at restoring similar functional activities. *Second*, we had no other way to define compliance than whether a patient had completed treatment or prematurely withdrew, whereas obviously more sophisticated ways are possible (e.g. evaluating the cooperation of a patient, or the quality of a patients homework). Nevertheless, since both groups considered their treatment equally credible, and because an equal number of patients prematurely terminated treatment, we have no reason to assume that the groups adhered differentially to their treatment.

In conclusion, this study was unable to demonstrate that EXP is a more effective treatment than GA for CLBP patients, irrespective of their level of pain-related fear. The mechanisms through which EXP induces change relative to GA might be the reduction of perceived harmfulness of activities and pain catastrophizing cognitions. In view of possible insufficient statistical power, the borderline significant difference between treatment conditions in favour of EXP calls for a replication of this trial with a larger sample of participants. Still, the fact that in the EXP condition approximately half the patients reported clinically relevant improvements, and the fact that EXP was not inferior to GA on any of our outcomes, renders EXP a worthwhile treatment for CLBP patients.

**ACKNOWLEDGEMENTS**

This study was supported by Grant nr. 1436.0002 of the Netherlands Organisation for Health Research and Development (ZonMw). We are grateful to Sita van Riet and Magda van Duinen for their great work with the data collection, to Rosanne Janssen for her help with the electronic data collection, and to Saskia Nijst for treating the majority of patients and additional research assistance. We are also thankful to all physicians referring patients to the study, to the patients willing to participate, and to the many therapists and logistical co-workers from the treatment centres of Maastricht, Hoensbroek, Sittard and Eindhoven, contributing to the successful completion of the study. Finally, we want to thank Jeanine Verbunt, Henk Seelen, Ivan Huijnen and Eric Bousema for their advice in the application and processing of the electronic activity monitors, Sarah Holla and Dominique Gielen for their evaluations of protocol adherence and contamination, and Stephen Morley, Steven Linton, and Sandra Beurskens for their helpful advice.

**Appendix A**

Dutch translation and modification of the credibility-expectancy questionnaire for patients.

**Instructie**

We willen u vragen om op de volgende bladzijden aan te geven hoe sterk u op dit moment gelooft dat de behandeling die u gaat volgen u zal helpen in: het leren uitvoeren van dagelijkse activiteiten, in het verminderen van beperkingen ondanks de pijn, en in het verminderen van de pijn.

De vragenlijst bestaat uit twee delen. In het eerste deel is het de bedoeling dat u aangeeft hoe u denkt over de behandeling *in het algemeen*. In het tweede deel is het de bedoeling dat u aangeeft hoe u denkt over de behandeling en het mogelijke succes ervan *voor uzelf*.

Beantwoord alstublieft de volgende vragen.

**Deel 1**

In dit eerste deel is het de bedoeling dat u aangeeft hoe u denkt over de behandeling ***in het algemeen***.

Omcirkel bij elke uitspraak het cijfer dat het beste bij uw antwoord past.

1. Hoe geloofwaardig lijkt op dit moment de behandeling die u aangeboden wordt?

1	2	3	4	5	6	7	8	9
helemaal niet				een beetje				heel erg

2. a. Hoe succesvol denkt u op dit moment dat de behandeling zal zijn in het leren uitvoeren van dagelijkse activiteiten ondanks de pijn?

1	2	3	4	5	6	7	8	9
helemaal niet				een beetje				heel erg

- b. Hoe succesvol denkt u op dit moment dat de behandeling zal zijn in het verminderen van beperkingen ondanks de pijn?

1	2	3	4	5	6	7	8	9
helemaal niet				een beetje				heel erg

- c. Hoe succesvol denkt u op dit moment dat de behandeling zal zijn in het verminderen van de pijn?

1	2	3	4	5	6	7	8	9
helemaal niet				een beetje				heel erg

3. Hoe zelfverzekerd zou u zijn in het aanraden van deze behandeling aan een vriend(in) met dezelfde klachten als u?

1	2	3	4	5	6	7	8	9
helemaal niet				een beetje				heel erg



Deel 2

Voor dit tweede onderdeel is het de bedoeling dat u aangeeft hoe u denkt over de behandeling en het mogelijke succes ervan **voor uzelf**.

Omcirkel bij elke uitspraak het cijfer dat het beste bij uw antwoord past.

4.

a.

Hoeveel verwacht u dat de behandeling **u** zal helpen in het leren uitvoeren van dagelijkse activiteiten ondanks de pijn?

1

2

3

4

5

6

7

8

9

helemaal niet

een beetje

heel erg

b.

Hoeveel verwacht u dat de behandeling **u** zal helpen in het verminderen van uw beperkingen ondanks de pijn?

1

2

3

4

5

6

7

8

9

helemaal niet

een beetje

heel erg

c.

Hoeveel verwacht u dat de behandeling **u** zal helpen in het verminderen van uw pijn?

1

2

3

4

5

6

7

8

9

helemaal niet

een beetje

heel erg
5.

a.

Hoeveel verbetering in **uw** activiteitsniveau verwacht u aan het einde van de behandelperiode?

1

2

3

4

5

6

7

8

9

helemaal niet

een beetje

heel erg

b.

Hoeveel verbetering in **uw** beperkingen verwacht u aan het einde van de behandelperiode?

1

2

3

4

5

6

7

8

9

helemaal niet

een beetje

heel erg

c.

Hoeveel verbetering in **uw** pijn verwacht u aan het einde van de behandelperiode?

1

2

3

4

5

6

7

8

9

helemaal niet

een beetje

heel erg

**Appendix B**

Dutch translation and modification of the credibility-expectancy questionnaire: the credibility scale for therapists

**Instructie**

We willen u vragen om op de volgende bladzijden aan te geven hoe sterk u op dit moment gelooft dat de behandelingen, die u in het kader van BackActive aan patiënten met chronisch lage rugpijn en bewegingsvrees gaat aanbieden, hun zal helpen in: het leren uitvoeren van dagelijkse activiteiten, in het verminderen van beperkingen ondanks de pijn, en in het verminderen van de pijn.

De vragenlijst bestaat uit twee delen. In het eerste deel is het de bedoeling dat u aangeeft hoe u denkt over de exposure in vivo behandeling. In het tweede deel is het de bedoeling dat u aangeeft hoe u denkt over de graded activity behandeling.

Beantwoord alstublieft de vragen op de volgende bladzijden.

Deel 1

In dit eerste deel is het de bedoeling dat u aangeeft hoe u denkt over de **exposure in vivo** behandeling voor patiënten met aspecifieke chronische lage rugpijn en bewegingsvrees (BackActive patiënten).

1.

Hoe geloofwaardig is exposure in vivo voor u op dit moment?

1

2

3

4

5

6

7

8

9

helemaal niet

een beetje

heel erg
2.

a

Hoe succesvol denkt u op dit moment dat exposure in vivo zal zijn in het leren uitvoeren van dagelijkse activiteiten ondanks de pijn?

1

2

3

4

5

6

7

8

9

helemaal niet

een beetje

heel erg
- b

Hoe succesvol denkt u op dit moment dat exposure in vivo zal zijn in het verminderen van beperkingen ondanks de pijn?

1

2

3

4

5

6

7

8

9

helemaal niet

een beetje

heel erg
- c

Hoe succesvol denkt u op dit moment dat exposure in vivo zal zijn in het verminderen van de pijn?

1

2

3

4

5

6

7

8

9

helemaal niet

een beetje

heel erg
3.

Hoe zelfverzekerd zou u zijn in het aanraden van exposure in vivo aan een vriend(in) met dezelfde klachten als deze patiëntgroep (aspecifieke chronisch lage rugpijn met bewegingsvrees)?

1

2

3

4

5

6

7

8

9

helemaal niet

een beetje

heel erg
4.

Hoe sterk gaat uw voorkeur uit naar exposure in vivo als uzelf dezelfde klachten zou krijgen als de BackActive patiënten?

1

2

3

4

5

6

7

8

9

helemaal niet

een beetje

heel erg
5.

Hoe sterk gaat uw voorkeur uit naar het geven van exposure in vivo boven graded activity bij een BackActive patiënt?

1

2

3

4

5

6

7

8

9

helemaal niet

een beetje

heel erg
6.

Hoe zelfverzekerd zou u zijn in het aanraden van exposure in vivo aan een collega, die werkt met dezelfde patiëntgroep als BackActive, en die nog niet bekend is met deze behandeling?

1

2

3

4

5

6

7

8

9

helemaal niet

een beetje

heel erg

**Deel 2**

In dit tweede deel is het de bedoeling dat u aangeeft hoe u denkt over de **graded activity** behandeling voor patiënten met aspecifieke chronische lage rugpijn en bewegingsvrees (BackActive patiënten).

1. Hoe geloofwaardig is graded activity voor u op dit moment?
 

1	2	3	4	5	6	7	8	9
helemaal niet				een beetje				heel erg
  
2. a. Hoe succesvol denkt u op dit moment dat graded activity zal zijn in het leren uitvoeren van dagelijkse activiteiten ondanks de pijn?
 

1	2	3	4	5	6	7	8	9
helemaal niet				een beetje				heel erg
  
- b. Hoe succesvol denkt u op dit moment dat graded activity zal zijn in het verminderen van beperkingen ondanks de pijn?
 

1	2	3	4	5	6	7	8	9
helemaal niet				een beetje				heel erg
  
- c. Hoe succesvol denkt u op dit moment dat graded activity zal zijn in het verminderen van de pijn?
 

1	2	3	4	5	6	7	8	9
helemaal niet				een beetje				heel erg
  
3. Hoe zelfverzekerd zou u zijn in het aanraden van graded activity aan een vriend(in) met dezelfde klachten als deze patiëntgroep (aspecifieke chronisch lage rugpijn met bewegingsvrees)?
 

1	2	3	4	5	6	7	8	9
helemaal niet				een beetje				heel erg
  
4. Hoe sterk gaat uw voorkeur uit naar graded activity als uzelf dezelfde klachten zou krijgen als de BackActive patiënten?
 

1	2	3	4	5	6	7	8	9
helemaal niet				een beetje				heel erg
  
5. Hoe sterk gaat uw voorkeur uit naar het geven van graded activity boven exposure in vivo bij een BackActive patiënt?
 

1	2	3	4	5	6	7	8	9
helemaal niet				een beetje				heel erg
  
6. Hoe zelfverzekerd zou u zijn in het aanraden van graded activity aan een collega, die werkt met dezelfde patiëntgroep als BackActive, en die nog niet bekend is met deze behandeling?
 

1	2	3	4	5	6	7	8	9
helemaal niet				een beetje				heel erg

## Appendix C

### The mixed model for testing treatment effects on the outcomes

The following mixed effect regression model was applied to each outcome variable  $Y$  as recorded on time point  $t$  for patient  $i$ :

$$Y_{it} = \beta_0 + \beta_1 \text{group} + \beta_2 \text{cov} + \beta_3 \text{post} + \beta_4 \text{foll} + \beta_5 \text{group} \times \text{post} \\ + \beta_6 \text{group} \times \text{foll} + \beta_7 \text{cov} \times \text{post} + \beta_8 \text{cov} \times \text{foll} + e_{it}$$

where:

- $t = 1, 2$  or  $3$  for baseline, post-test or follow-up 1, respectively;
- $\text{group} = 1$  for patients assigned to EXP treatment, and  $0$  for patients assigned to GA;
- $\text{cov}$  = covariates, which are enumerated in section 'treatment outcome' in statistical analyses, with quantitative covariates being standardized (mean =  $0$ , sd =  $1$ ), and others remaining dichotomous;
- $\text{post} = 1$  if  $t = 2$  and  $0$  else, that is, post is a dummy indicator for the posttest time point;
- $\text{foll} = 1$  if  $t = 3$  and  $0$  else, that is, foll is a dummy indicator for the follow-up 1 time point;
- $e_{it}$  = a random effect of patient  $i$  at time point  $t$ .

This model must be interpreted as follows:

- $\beta_0$  = the outcome mean at baseline (post=0, foll=0) in the GA group (group =  $0$ );
- $\beta_1$  = the mean baseline difference between both groups (EXP - GA), which must be zero apart from sampling error due to the randomized treatment assignment, and may therefore be deleted from the model in order to gain power (for details, see: [Laird, 1990 #1188; van Breukelen, 2006 #1103];
- $\beta_2$  = the association between the covariate and the outcome
- $\beta_3$  = the mean change from baseline to posttest within patients who score  $0$  on all predictors included in the final model (e.g. group = GA, gender = female, mean score on covariates);
- $\beta_4$  = the mean change from baseline to follow-up 1 within patients who score  $0$  on all predictors included in the final model (e.g. group = GA, gender = female, mean score on covariates);
- $\beta_5$  = the group difference (EXP-GA) in mean change from baseline to posttest, which is also the group difference at posttest, since there is no group difference at baseline;
- $\beta_6$  = the group difference (EXP-GA) in mean change from baseline to follow-up 1, which is also the group difference at follow-up 1, since there is no group difference at baseline;
- $\beta_7$  = the effect of a specific covariate on the change from baseline to posttest in both treatment conditions;
- $\beta_8$  = the effect of a specific covariate on the change from baseline to follow-up 1 in both treatment conditions;

This model was extended with group by covariate by time interactions which were again dropped if not significant according to a likelihood ratio test. Group by covariate effects at baseline were not expected or found anyway, for the same reason as the group main effect at baseline, that is, due to the randomized treatment assignment.

The null hypothesis of no difference between GEXP and GA implies that  $\beta_5 = \beta_6 = 0$ . This null hypothesis was tested against the alternative of a difference between the treatments at post-test or follow-up1, with a likelihood ratio test with  $df = 2$ .

The null hypothesis of an equal difference at both time points, post-test and follow-up1, implies that  $\beta_5 = \beta_6 \neq 0$ . This null hypothesis was tested against the alternative that  $\beta_5 \neq \beta_6$ , both  $\neq 0$ , with a likelihood ratio test with  $df = 1$ .

The 3 random student effects  $e_{1ij}$ ,  $e_{2ij}$ ,  $e_{3ij}$  were assumed to be multivariate normally distributed with unspecified covariance structure, which is the same assumption as in a repeated measures ANOVA using the so-called multivariate method of analysis, and more general than compound symmetry at the expense of a few degrees of freedom.

## Chapter 8

### **Assessing protocol adherence and contamination in treatment outcome studies: A methodological illustration from behavioural medicine**

**Maaïke Leeuw, Mariëlle E.J.B. Goossens,  
Henrica C.W. de Vet, Johan W.S. Vlaeyen**

**This chapter is accepted for publication, pending revision:**

Leeuw M, Goossens MEJB, de Vet HCW, Vlaeyen JWS. Assessing protocol adherence and contamination in treatment outcome studies: A methodological illustration from behavioural medicine. *Journal of Clinical Epidemiology*, accepted for publication pending revision.

**ABSTRACT**

*Objective:* Treatment outcome studies ought to assess the fidelity of their treatments, including protocol adherence, but practical guidelines and examples for this are lacking. Based on general recommendations in available literature, this study proposes and illustrates the design and application of a Method of assessing Protocol Adherence and Contamination (MPAC) in a behavioural medicine trial evaluating two treatments for chronic low back pain.

*Study Design and Setting:* In designing the MPAC, two experts identified several feasible treatment elements. Agreement between the experts in classifying these elements into five categories (essential-and-unique, essential-but-not-unique, unique-but-not-essential, compatible, prohibited) was assessed, as well as whether trained therapists concurred with their categorizations. In applying the MPAC, treatment recordings were evaluated by two independent raters, who coded the (non)-occurrence of the MPAC elements and who categorized each session as belonging to one of both treatments.

*Results:* The MPAC's content validity was supported by adequate agreement between the experts' classifications of the treatment elements. These classifications were sufficiently corroborated by the trained therapists. The MPAC's interrater reliability was good.

*Conclusion:* Comprehensive illustrations of designing and applying an MPAC may encourage the assessment of protocol adherence as a partial reflection of treatment fidelity in forthcoming treatment outcome studies.

## INTRODUCTION

Treatment fidelity, which may also be denoted as integrity, is defined as the extent to which a treatment is carried out as intended (Yeaton and Sechrest, 1981). Treatment outcome studies ought to assess the fidelity of the treatments delivered to ensure honest and genuine comparisons (Yeaton and Sechrest, 1981; Peterson et al., 1982; Moncher and Prinz, 1991; Mayo-Wilson, 2007). Cook and Campbell (1979) even asserted that ‘measures of the exact nature of the treatment in *all* treatment and control groups are absolutely vital in any experiment’ (p.59). Especially psychological interventions may be at higher risk of compromised treatment fidelity, since these are generally more complex and extensive (Yeaton and Sechrest, 1981; Peterson et al., 1982; Moncher and Prinz, 1991; Pereplechikova and Kazdin, 2005). The absence of treatment fidelity checks can seriously obscure the conclusions about treatment effectiveness: in case a treatment is found to be effective, this may be due to unknown contaminants, whereas in case of an ineffective treatment one cannot rule out the possibility that this is because the treatment was inadequately carried out (Moncher and Prinz, 1991; Gresham et al., 1993; Waltz et al., 1993). Besides jeopardizing the *internal* validity and statistical power, insufficient treatment fidelity may also compromise the *external* validity of treatment outcome studies (Cook and Campbell, 1979; Moncher and Prinz, 1991; Gresham et al., 1993; Bellg et al., 2004). Notwithstanding its importance, several reviews demonstrated that the majority of studies have not verified whether the fidelity of their treatments was adequate (Peterson et al., 1982; Moncher and Prinz, 1991; Gresham et al., 1993; Lichstein et al., 1994; Borrelli et al., 2005).

Various opinions exist on the number of components that treatment fidelity consists of. According to Pereplechikova and Kazdin (2005), treatment fidelity consists of three components. These are *protocol adherence*, referring to the degree to which specific treatment procedures are utilized by the therapists during actual delivery of treatment (Waltz et al., 1993), *competence*, which is the skillfulness of the therapists delivering the treatment (Waltz et al., 1993), and *differentiation*, signifying whether the therapies differ from other treatments on several critical dimensions, and especially with respect to the manipulation of the variable of interest (Kazdin, 1986). Others added two other aspects of treatment fidelity to these, which are the *treatment receipt* by the patient, e.g. if the patient understands and is able to utilize the treatment skills, and a patient’s *enactment* upon treatment, e.g. whether the patient is able to actually implement the learned behaviour in daily life (Lichstein et al., 1994; Bellg et al., 2004; Borrelli et al., 2005). The difficulty in assessing treatment fidelity may depend on which aspect of this is measured. Of all aspects, protocol adherence might be the least complex and time-consuming to assess. We therefore plead that treatment outcome studies at least determine protocol adherence, although bearing in mind that this is only part of the full comprehension of treatment fidelity. Besides assuring the occurrence of essential components, the assessment of protocol adherence also includes a verification of the non-occurrence of protocol deviations that are not supposed to happen, e.g. the absence of treatment contamination (Peterson et al., 1982; Moncher and Prinz, 1991; Waltz et al., 1993; Borrelli et al., 2005).

Depending on the purpose, the manner in which protocol adherence can be assessed may differ. First, protocol adherence and contamination can be evaluated



concurrently and repeatedly during the trial to improve these by providing feedback to the treatment agents. For this, it is necessary to reflect for *each individual treatment element* whether it is adequately carried out in case of required elements, and whether it is sufficiently absent in case of prohibited ones. By this means, protocol adherence can be optimized, and treatment contamination reduced, specifically with respect to those elements for which this appears to be relevant (Cook and Campbell, 1979; Moncher and Prinz, 1991; Gresham et al., 1993; Bellg et al., 2004; Pereplechikova and Kazdin, 2005). Second, this assessment can be carried out after completion of the trial in order to determine whether treatment outcome comparisons are fair in the sense that both treatments were equally and adequately carried out according to their protocols without being contaminated with proscribed treatment components. For this, it may be sufficient to appraise to *overall protocol adherence and contamination* for both treatments, that is whether on average the obliged treatment elements are addressed whereas prohibited components are not, and to ensure that in both treatment conditions the therapists equally adhered to the treatment manuals.

Even though some papers present general recommendations for the assessment of protocol adherence as well as other aspects of treatment fidelity (Gresham et al., 1993; Waltz et al., 1993; Bellg et al., 2004; Pereplechikova and Kazdin, 2005), these remain rather indefinite, so researchers have to design their own specific method of doing this. In addition, the studies indicating to have assessed treatment fidelity did not describe in detail how this was evaluated, probably due to space limitations imposed by most journals (Moncher and Prinz, 1991; Gresham et al., 1993). Furthermore, since the assessment of protocol adherence frequently requires subjective judgments, the authentication of the reliability and validity of the assessment method is essential (Peterson et al., 1982). The aim of this study is to propose a Method of assessing Protocol Adherence and Contamination (MPAC) based on crucial elements presented in the literature<sup>5</sup>. In this paper, we will apply this method to a trial in the field of behavioural medicine, to verify whether fair effectiveness comparisons were made between two multidisciplinary behavioural treatments. By this, the method's feasibility will be tested, and it will be comprehensively illustrated how the method was designed and applied. Although the exact specifications of such a method are determined by the contents of the treatments of interest (Pereplechikova and Kazdin, 2005), the detailed description of our method will aid other researchers in constructing their own assessment tool. First, some relevant background information about the interventions, and our efforts to enhance treatment fidelity in this study, are outlined.

## BACKGROUND INFORMATION

There is accumulating evidence that in chronic low back pain, fear of pain is more disabling than pain itself (Vlaeyen and Linton, 2000; Leeuw et al., 2007a). Since pain-related fear has been found to contribute to the development and maintenance of chronic low back pain and disability, it has recently been chosen as an important target for intervention when aiming to reduce disability in these patients. Preliminary support was found for exposure in vivo (EXP) as an effective intervention in diminishing

---

<sup>5</sup> We mainly relied on recommendations presented in existing literature. For these, we refer to the relevant source. When we encountered feasibility issues during designing and applying the MPAC for which we did not encounter any information in the literature, we added some personal undertakings to these. These will be recognized by the absence of a reference.

disability by the systematic reduction of pain-related fear (e.g. Vlaeyen et al., 2002b; Boersma et al., 2004; de Jong et al., 2005b). In a multi-centre randomized controlled trial (ISRCTN88087718), we compared the effectiveness of EXP with operant graded activity (GA) in chronic low back pain patients. EXP systematically aims to reduce pain-related fear by gradually exposing fearful patients to threatening previously avoided activities. Due to this, they discover that the feared consequences are catastrophic misinterpretations, and their fear and disability may diminish consequently (Vlaeyen et al., 2002c; Vlaeyen et al., 2004a). In GA, the main focus lies on operant mechanisms of pain behaviour: healthy behaviour is encouraged, and illness behaviour discouraged, by learning patients to gradually increase their activity levels according to a time contingent treatment plan, and by providing positive reinforcers when the patients accomplish these predefined activity quota (Sanders, 2002).

Both treatments were carried out in four treatment centres. Different combinations of psychologists with physiotherapists or occupational therapists resulted in 19 therapist-teams. All therapists had at least half a year experience in treating chronic pain patients in a rehabilitation setting. The teams performed both treatments to ensure that the general qualities of the therapists were evenly distributed across conditions, although the psychologists were less involved in GA than in EXP. Both EXP and GA were structuralized and consisted of approximately 16 and 26 sessions respectively. The treatment sessions of EXP were carried out by both a psychologist and a physiotherapist or occupational therapist, whereas those of GA were carried out solely by a physiotherapist or occupational therapist, except for two psychological sessions.

We found that EXP, despite its superior ability in reducing pain-related fear, was equally effective as GA in reducing functional disability and main complaints, although the difference between treatment conditions almost reached statistical significance favouring EXP.

We undertook several efforts to enhance the fidelity of the treatments delivered, of which several are also described by Moncher & Prinz (1991) and Pereplechikova and Kazdin (2005). First, we tried to optimize *protocol adherence* of the therapists by explaining the treatment rationales and their effectiveness to them. Second, the therapists were provided with detailed treatment manuals that gave specific instructions for each session, and that provided guidelines for handling commonly encountered problems (Luborsky and DeRubeis, 1984). Third, therapists had to record their therapy sessions and were explicitly told that these tapes would be used to verify whether protocol adherence was sufficient and treatment contamination absent. *Therapeutic competence* was first encouraged by selecting therapists who were motivated to carry out both treatments, and to restrict themselves to the treatment manuals. This was further promoted by comprehensive training of the therapists in both treatments by specialists of the treatment protocols, during which plenty opportunity for practice and role-playing was available. Three times a year, collective supervision sessions were provided. In between supervision sessions, therapists were able to individually consult several experts of the treatment manuals (M.L., J.V., and M.G.). *Differentiation* between treatments was promoted by the detailed treatment

protocols, in addition to extensive outlining of the theoretical and practical differences of the treatments to the therapists.

We wanted to verify whether effectiveness comparisons between EXP and GA were fair. For this, we designed an MPAC that assessed these aspects for the evaluated sessions in general, and applied this after completion of the trial. In doing so, we also focussed on the verification of the content validity and reliability of this method (Peterson et al., 1982; Pereplechikova and Kazdin, 2005). The subsequent steps that we undertook in the design and application of the MPAC are summarized in Table 1.

## **DESIGNING THE MPAC**

An MPAC needs to be capable of verifying whether important treatment elements are actually addressed (maximal protocol adherence), and whereas proscribed elements do not occur (minimal contamination), during delivery of treatment (Peterson et al., 1982; Moncher and Prinz, 1991; Waltz et al., 1993; Borrelli et al., 2005; Pereplechikova and Kazdin, 2005). In case of comparing two interventions, there most likely will be some treatment components that apply to both treatment conditions, but also those that are unique for a particular treatment and that are therefore proscribed in the contrasting condition. The method has to represent a range of feasible elements defined by the treatment manual of each treatment, enabling for comparing the treatments on all of the discriminating as well as overlapping components (Waltz et al., 1993). We therefore developed a single MPAC that aimed to determine protocol adherence and treatment contamination for EXP and GA simultaneously. Evaluations of EXP sessions are expected to be reflected in positive scores on the unique and general required treatment elements of EXP, and in negative scores on those unique to GA (and thus proscribed for EXP), whereas the opposite applies for the GA sessions. The design of the MPAC consists of five steps.

### **Step 1. Dividing the treatment into phases**

In the literature, we did not encounter any information or recommendations regarding the application of the MPAC with respect to different stages of treatment. Nevertheless, this is necessary when the content of a treatment substantially differs between sessions, and thus when differential presence/absence of treatment elements is required for different phases of treatment. For example, the fact that a behavioural experiment is not yet performed during the explanation of the treatment rationale does not imply that protocol adherence is insufficient, but rather that this element is only essential later on during treatment. Since each phase needs its own compilation of treatment components, the number of treatment phases has to be minimized.

In our study, we divided EXP and GA in three distinct phases: the preparation phase, that consisted of the establishment of a fear hierarchy in EXP, and of a patient's baseline activity level in GA; The educational phase, during which in both treatments the treatment rationale was explained to the patient, and; The treatment phase, that consisted of behavioural experiments and exposure in vivo in EXP, and of positive reinforcement of time contingent activity increases in GA.

**Table 1**

List of subsequent steps that we undertook in designing and applying a Method to assess Protocol Adherence and treatment Contamination (MPAC)

## DESIGNING THE MPAC

### 1. Dividing the treatment into phases

We divided our treatments into three conceptually different phases.

### 2. Identification of possible treatment elements

Two experts of both treatment protocols identified possible treatment elements, including elements of varying importance as well as prohibited ones, per phase for both treatment conditions.

### 3. Categorization of treatment elements

The experts who previously identified possible treatment elements independently categorized these elements into five categories for each phase of each treatment separately: (1) Essential-and-unique; (2) Essential-but-not-unique; (3) Unique-but-not-essential; (4) Compatible, but not essential and not unique; (5) Prohibited. The same was requested from the therapists that participated in the trial and who were trained in both treatment protocols.

### 4. Establishing the content validity of the MPAC

The two experts agreed in their categorization in 80% of the items (Cohen's kappa = .73), which was higher than the predetermined criteria for sufficient agreement. For the remaining items, they reached consensus after single consultation. The mean agreement of the therapists with this final categorization of the experts was 73%, which was again higher than the predetermined criteria for adequate agreement. These results support the content validity of the MPAC. Based on these findings, the final categorization of the experts was taken as the point of reference for determining whether treatment elements were required to occur (essential-and-unique, essential-but-not-unique), allowed to occur (unique-but-not-essential, compatible), or proscribed (prohibited) to occur.

### 5. Making a scoring form for the MPAC

The treatment elements were listed in random order in the final MPAC separate for each phase, without revealing to the users of this method whether these elements were required, allowed, or prohibited to occur. For each item, the (non)occurrence during treatment delivery could be indicated. The MPAC also included the question 'which treatment do you believe that the current session belongs to?' together with response choices listing the treatment conditions.

## APPLYING THE MPAC: EVALUATING TREATMENT DELIVERY

### 1. Recording of treatment sessions

Out of approximately 1275 sessions, only 265 were recorded, from which a random, but representative, selection of 30 sessions was drawn.

### 2. Selecting raters

Two undergraduate students who were not involved in the trial were trained and paid to rate the selected recordings with the MPAC.

### 3. Scoring of the treatment sessions

The raters interrupted their listening to the recordings every five minutes, and filled in the MPAC for the relevant elements in between times. They did not rate the same session simultaneously.

### 4. Calculation of interrater reliability

The agreement found between both raters, as represented by Cohen's kappa, was .72 for the specific treatment elements and .87 for the categorization of treatment condition, indicating good interrater reliability for the MPAC.

### 5. Calculation of protocol adherence

By dividing the number of observed required treatment elements (essential-and-unique, and essential-but-not-unique) by the maximum possible number of these elements, we found that 72% of these elements were addressed during all treatment sessions, which exceeded the predetermined criterion of 70% for sufficient protocol adherence. Furthermore, in 97% of the cases the raters categorized the session to the correct treatment, which exceeded the predetermined criterion of 90% for adequate protocol adherence.

### 6. Calculation of treatment contamination

By dividing the number of observed prohibited treatment elements by the maximum possible number of these elements, we found that 4% of these elements were addressed during all treatment sessions, which fulfilled the predetermined criterion for sufficiently low treatment contamination.

## Step 2. Identification of possible treatment elements

An MPAC includes elements that could occur during treatment, comprising elements that are required, allowed, or not allowed during the treatments of interest (Waltz et al., 1993; Borrelli et al., 2005; Pereplechikova and Kazdin, 2005). In case of protocolized interventions, the identification of these elements should be rather straightforward. However, this can be complicated for trials where treatment-as-usual is the control condition, for which in the absence of guidelines it is intricate to identify

elements that are required or prohibited to occur. The prohibited treatment elements may not only consist of elements that are proscribed for the treatments of interest in general (e.g. giving a sole biomedical explanation of low back pain within a bio-psycho-social intervention), but also of those that are unique to the contrasting treatment (e.g. given that ‘integrating a behavioural experiment’ is clearly a unique EXP element, this is prohibited in GA). These elements have to be determined separately for each of the previously identified treatment phases. Obviously, these elements can best be identified by persons with profound knowledge of the treatments of interest. Detailed definitions of these elements may be useful to resolve any ambiguity about the denotation of a treatment component.

In our study, we decided to have two experts of both treatment protocols (J.V. and M.L.) identify various possible treatment elements for EXP and GA separately for the three treatment phases. The elements of GA and EXP were then intermixed and listed in a random order. This resulted in 16 items for the preparatory phase (Table 2), 24 items for the educational phase (Table 3), and 20 items for the treatment phase (Table 4). Unfortunately, we did not provide detailed definitions of these treatment elements at this stage. So far, the experts refrained from denoting the importance of the (non)-occurrence of these elements.

**Table 2**  
The categorization and percentage of occurrences observed of the specific elements of the preparatory phase of the MPAC displayed separately for EXP and GA (n = 5 recorded sessions evaluated per treatment condition)

No.	Specific elements of the preparatory phase	Category			
		EXP (% present)		GA (% present)	
1	The patient's baseline level of activities is being determined	P	(0)	EU	(50)
2	There is good teamwork between the therapists and patient	E	(100)	E	(100)
3	The patient's concern/fear with regard to activities is being discussed	EU	(30)	P	(10)
4	Enquiries are made about the patient's feelings and mood	C	(10)	C	(20)
5	The therapists go into possible medical causes of the symptoms	P	(0)	P	(0)
6	The patient assesses the level of perceived threat value of daily activities	EU	(100)	P	(0)
7	The patient performs activities in a pain-contingent manner	P	(0)	EU	(100)
8	The aim of the current session is being explained to the patient	E	(70)	E	(90)
9	The therapists respond understandingly to the problems expressed by the patient	E	(100)	E	(100)
10	Photographs of daily activities are being used	EU	(100)	P	(0)
11	Supportive activities are being determined or carried out	P	(0)	U	(80)
12	The performance of activities is being recorded	P	(0)	EU	(90)
13	The therapists use medical terminology (e.g. diagnostic labels)	P	(0)	P	(0)
14	A hierarchy is being developed based on the threat value of daily activities	EU	(100)	P	(0)
15	Important functional activities are being determined or performed	P	(30)	EU	(100)
16	The therapists express fear/concern with regard to pain or activities	P	(0)	P	(0)

EU = Essential-and-unique; E = Essential-but-not-unique; U = Unique-but-not-essential; C = Compatible but not essential and not unique; P = Prohibited

**Table 3**

The categorization and percentage of occurrences observed of the specific elements of the *educational phase* of the MPAC displayed separately for EXP and GA (n = 4 recorded sessions evaluated per treatment condition)

No.	Specific elements of the educational phase	Category			
		EXP (% present)		GA (% present)	
1	It is emphasized that in chronic pain no clear relationship exists between pain and injury	E	(0)	E	(50)
2	The circular model of pain-inactivity-pain is being explained	C	(13)	EU	(75)
3	The patient's concern/fear with regard to activities is being discussed	EU	(63)	P	(0)
4	There is good teamwork between the therapists and patient	E	(100)	E	(88)
5	The aim of the current session is being explained to the patient.	E	(88)	E	(75)
6	Examples of positive reinforcements are being identified.	P	(0)	EU	(75)
7	A fluctuating pain-contingent activity pattern is being discussed	P	(13)	EU	(100)
8	There is an understanding attitude with regard to the patient's behaviour	C	(100)	C	(100)
9	A biomedical approach to pain is being discouraged	E	(0)	E	(50)
10	It is explained that the treatment is aimed at ending inactivity	P	(13)	EU	(38)
11	It is emphasized that pain reduction is not a therapy goal	E	(88)	E	(75)
12	The patient is being actively involved in the explanation of the therapy	E	(100)	E	(100)
13	It is explained that the treatment is aimed at verifying examining cognitions	EU	(75)	P	(0)
14	A bio-psycho-social approach to pain is being explained	E	(25)	E	(0)
15	It is emphasized that all activities are possible	E	(38)	E	(13)
16	The therapists use medical terminology (e.g. diagnostic labels)	P	(0)	P	(0)
17	The drawbacks of inactivity are being explained	E	(25)	E	(63)
18	The therapists go into possible medical causes of the symptoms.	P	(0)	P	(25)
19	The patient's motivation for the therapy is being checked	C	(100)	C	(63)
20	Enquiries are made about the patient's feelings and mood	C	(38)	C	(13)
21	The circular model pain - pain cognitions – avoidance - pain is being explained	EU	(88)	P	(0)
22	The therapists respond understandingly to the problems expressed by the patient.	E	(100)	E	(100)
23	It is explained that the aim of the therapy is an increase in activity level	E	(50)	E	(75)
24	The therapists express fear/concern with regard to pain or activities	P	(0)	P	(0)

EU = Essential-and-unique; E = Essential-but-not-unique; U = Unique-but-not-essential; C = Compatible but not essential and not unique; P = Prohibited

**Table 4**

The categorization and percentage of occurrences observed of the specific elements of the *treatment phase* of the MPAC displayed separately for EXP and GA (n = 6 recorded sessions evaluated per treatment condition)

No.	Specific elements of the treatment phase	Category			
		EXP (% present)		GA (% present)	
1	Homework is being assigned	E	(58)	E	(25)
2	The level of activities is being increased in a time-contingent manner	P	(8)	EU	(75)
3	The patient's concern/fear with regard to activities is being discussed	EU	(92)	P	(0)
4	The aim of the current session is being explained to the patient.	E	(0)	E	(33)
5	A catastrophizing cognition is being identified	EU	(83)	P	(0)
6	There is good teamwork between the therapists and patient	E	(92)	E	(100)
7	The therapists use medical terminology (e.g. diagnostic labels)	P	(0)	P	(0)
8	A behavioural experiment is being performed	EU	(75)	P	(0)
9	The therapists respond understandingly to the problems expressed by the patient.	E	(100)	E	(100)
10	A catastrophizing cognition is being evaluated	U	(67)	P	(0)
11	Activities from the hierarchy or based on threat value are being performed.	EU	(67)	P	(0)
12	Enquiries are made about the patient's feelings and mood	C	(33)	C	(50)
13	The therapists express fear/concern with regard to pain or activities	P	(0)	P	(8)
14	Clear agreements are made about the way in which an activity should be carried out (e.g. how often, how high the jumps should be, how to bend down)	EU	(42)	C	(17)
15	Homework is being evaluated	E	(67)	E	(58)
16	Activities are being carried out according to a time-contingent plan of treatment	P	(8)	EU	(75)
17	There is positive reinforcement of activity quotas that are met	P	(0)	EU	(8)
18	The therapists pay a lot of attention to the patient's pain behaviour	P	(50)	P	(8)
19	The performance of activities is being recorded	C	(8)	E	(67)
20	The therapists go into possible medical causes of the symptoms.	P	(0)	P	(0)

EU = Essential-and-unique; E = Essential-but-not-unique; U = Unique-but-not-essential; C = Compatible but not essential and not unique; P = Prohibited

**Step 3. Categorization of the treatment elements**

As recommended, each of these previously identified treatment elements has to be classified into one of the following categories: (1) essential-and-unique; (2) essential-but-not-unique; (3) compatible-but-not-essential-and-not-unique; (4) prohibited (Waltz et al., 1993; Pereplechikova and Kazdin, 2005), to which we added a fifth category, that is (5) unique-but-not-essential. Since it is important to evaluate the validity of such method (Pereplechikova and Kazdin, 2005), we suggest that two experts of the treatment protocols independently allocate these elements to the previously mentioned categories. Sufficient agreement between these experts in categorizing these elements can be conceived as support for the content validity of the MPAC.

In our study, the two experts who earlier identified the feasible treatment elements, now independently assigned these to the formerly mentioned five categories ranging from ‘essential-and-unique’ to ‘prohibited’ for the three phases for EXP and GA separately. So, different categorizations were collected for each phase of each treatment condition. The same procedure was requested of the therapists that participated, and were trained, in the study. Of 19 available therapists, 18 (7 psychologists and 11 physical or occupational therapists) completed this categorization.

#### **Step 4. Establishing the content validity of the MPAC**

Adequate agreement between experts in allocating the treatment elements to the same category of importance can be verified by calculating Cohen’s kappa. For the elements for which conformity is not found, agreement may either be achieved by consultation between these experts, or these may be removed from the MPAC.

In our study, the agreement between the experts in categorizing the treatment elements was adequate (Cohen’s kappa = .73). For example, the experts agreed that ‘a behavioural experiment is being performed’ and ‘the level of activities is being increased in a time-contingent manner’ are essential-and-unique elements for the treatment phase of EXP and GA respectively. For the other items, they reached consensus after a single consultation.

The final agreement of these experts was taken as a point of reference to which the categorization of the therapists was compared. We set an a priori level of agreement among experts and therapists at 70%. This may appear somewhat low, but we took into consideration that the majority of therapists did not gain much experience with EXP, whereas we requested them to categorize these elements not until at least nine months after finishing the trial, as well as the absence of clear definitions of the treatment elements. The mean agreement *per therapist* with the final categorization of the experts was 73% (SD = 8.15; EXP: M = 73.10, SD = 8.03; GA: M = 72.73, SD = 8.91). The mean agreement *per item* between the therapists and the final categorization of the experts was also 73%, though there was substantial variation between items (SD = 25.67; EXP: M = 72.80, SD = 25.46; GA: M = 72.80, SD = 26.10). For example, whereas the experts considered ‘Supportive activities are being determined or carried out’ as unique-but-not-essential to GA, none of the therapists agreed with this categorization. Notwithstanding this variability in agreement between experts and therapists, in general the agreement was sufficient to conclude that the content validity of the treatment elements of the MPAC was confirmed. Since the categorization of the experts was sufficiently corroborated by the therapists, their categorization was taken as the point of reference to determine which treatment elements were essential/essential-and-unique/unique-but-not-essential/compatible/prohibited, during the three phases for EXP and GA separately. The final categorizations of the treatment elements are displayed in Table 2 to 4<sup>6</sup>.

---

<sup>6</sup> With respect to the categorization of the experts it should be noted that, while unique elements actually should not be allowed in the contrasting treatment, the experts categorized two items to be unique to one treatment, but compatible to the other (item 2 of the educational phase, and item 14 of the treatment phase).



Step 5. Making a scoring form for the MPAC

When the categorization of the importance of the treatment elements is sufficiently agreed on, a scoring form for the MPAC can be constructed. The scoring form lists all of these previously determined treatment elements (Waltz et al., 1993; Borrelli et al., 2005; Pereplechikova and Kazdin, 2005), behind which it can be indicated whether it did or did not occur during treatment (Gresham, 2005; Pereplechikova and Kazdin, 2005). It may also comprise a question as to which treatment condition the administered treatment session belongs to (Borrelli et al., 2005).

In our study, per phase all treatment elements were listed in a random order without revealing whether these elements were essential/essential-and-unique/unique-but-not-essential/compatible/prohibited. After each treatment element, a dichotomous response choice was presented, so that its occurrence or non-occurrence during treatment delivery could be indicated. The MPAC also included the question ‘which treatment do you believe that the current session belongs to?’ together with response choices listing these treatment conditions (Figure 1).

**Figure 1**  
Example of a scoring form of an MPAC

Listen to the recording of the treatment session. Indicate for each of the specific treatment elements listed below whether these did or did not occur during this treatment session. Interrupt listening every 5 minutes and complete the relevant items in between times.

	Did occur	Did not occur
Specific treatment element 1	<input type="checkbox"/>	<input type="checkbox"/>
Specific treatment element 2	<input type="checkbox"/>	<input type="checkbox"/>
Specific treatment element 3	<input type="checkbox"/>	<input type="checkbox"/>
.....	<input type="checkbox"/>	<input type="checkbox"/>
.....	<input type="checkbox"/>	<input type="checkbox"/>
.....	<input type="checkbox"/>	<input type="checkbox"/>
.....	<input type="checkbox"/>	<input type="checkbox"/>
.....	<input type="checkbox"/>	<input type="checkbox"/>

Which treatment do you believe that the current session belongs to?

☐ Treatment A

☐ Treatment B

We added detailed definitions of the specific treatment elements to the MPAC, so that the elements were unambiguous to future users completing the MPAC. As can be seen in Table 2 to 4, there is considerable overlap between both treatment conditions on several of the treatment elements: in total, all essential-but-not-unique items were applicable to both treatments, in addition to five compatible and ten prohibited

elements. However, the MPAC also comprised treatment elements that were unique to EXP and thus proscribed for GA, and vice versa.

## **APPLYING THE MPAC: EVALUATING TREATMENT DELIVERY**

### **Step 1. Recording of treatment sessions**

Adequate representations of treatment delivery are best collected prospectively, that is during actual delivery of treatment. Audio recordings are suitable and preferred over self-report measures of patients or therapists, since these are objective reflections of actual treatment delivery (Moncher and Prinz, 1991; Bellg et al., 2004; Perepletchikova and Kazdin, 2005). Videotaping might even be more suitable, since this provides information about non-verbal communication that will not be available otherwise. Treatment sessions ought to be recorded as many as possible. Subsequently, a random selection can be drawn from these for actual assessment (Perepletchikova and Kazdin, 2005). The random selection ideally reflects the variation in sessions and patients, to optimize generalization of the findings throughout treatment and among patients (Moncher and Prinz, 1991; Gresham et al., 1993; Perepletchikova and Kazdin, 2005).

We provided each treatment centre with an MP3 recording device. Out of approximately 1275 performed treatment sessions, 265 (21%) were recorded. The median number of recorded sessions per therapist team was 5 (SD = 14.96, min. = 0, max. = 47). From these recordings, the part in which the therapists explicitly mentioned the treatment condition was removed (Borrelli et al., 2005). Subsequently, we drew a random selection of 30 sessions (11% of the available recordings) from these to be rated for protocol adherence and contamination, since statistically this number of sessions was minimally required to calculate interrater reliability. It was taken care of that for each team of therapists both an EXP and a GA session were selected if possible, that these sessions were derived from all three treatment phases, that a session was completely recorded, and that the quality of recording was sufficient to understand the conversation. Inadequate recordings were replaced by another randomly selected one until these criteria were met. Eventually, this resulted in respectively five, four, and six recordings for the preparatory, educational and treatment phase of EXP as well as GA.

### **Step 2. Selecting raters**

When intended to verify the interrater reliability of the MPAC, two raters are needed, who independently rate the selected treatments recordings with this measure. When these ratings are straightforward, undergraduate students with an education relevant for the treatments of interest (i.e. psychology or physical therapy student in case of rehabilitation treatments) may be capable of assessing the (non-)occurrence of treatment elements during treatment delivery. However, when the evaluation requires a certain expertise, experienced therapists are more suitable (Waltz et al., 1993). It is important to keep these raters blind with regard to the study hypotheses and as independent of the study as possible (Waltz et al., 1993; Borrelli et al., 2005; Perepletchikova and Kazdin, 2005). These raters have to be presented with sufficient information about the treatment condition(s), and it may be helpful to accustom them

to using the MPAC by practice (Moncher and Prinz, 1991; Perepletkhikova and Kazdin, 2005).

In our study, two undergraduate students in clinical psychology were trained as raters, since we believed that the assessment of the presence or absence of the treatment elements was fairly straightforward. These students had no previous involvement in the study and were kept blind with regard to the study hypotheses, except for their involvement in the assessment of interrater reliability. They were paid for the training as well as for the actual scoring of the sessions. They first gained knowledge of both treatments by reading the detailed treatment manuals. Then, they practiced three sessions together with one of the developers of the MPAC (M.L.) until sufficient understanding and agreement about the employment of the instrument was reached. After this, each student rated the 30 selected recordings of treatment sessions.

### **Step 3. Scoring of the treatment sessions**

We did not encounter any recommendations in the literature with regard to the scoring of the treatment sessions. To ensure adequate ratings, we instructed our raters to interrupt their listening to the recordings every five minutes, and to fill in the MPAC for the relevant elements in between times. Consultation between raters was prevented.

### **Step 4. Calculation of interrater reliability**

The interrater reliability between the raters (Peterson et al., 1982; Moncher and Prinz, 1991) can be determined by calculating Cohen's kappa, both for the specific treatment elements and the treatment categorization. Providing adequate interrater reliability, the mean of the two raters can be taken to calculate the mean proportion of essential or prohibited treatment elements.

In our study, the agreement found between both raters, as represented by Cohen's kappa, was .72 for the specific treatment elements and .87 for the categorization of treatment condition, indicating good interrater reliability for the MPAC.

### **Step 5. Calculation of protocol adherence**

A priori, it has to be determined to what degree required treatment elements should have occurred during treatment, as well as to what degree treatment sessions have to be correctly classified, to consider protocol adherence as adequate (Perepletkhikova and Kazdin, 2005). As far as we know, there are no clear guidelines available to determine whether protocol adherence is sufficient or not. Whereas Perepletkhikova and Kazdin (2005) propose that 'high integrity levels may be represented by 80-100% integrity, whereas low integrity condition may be represented by 50% integrity or less' (p. 377), it remains unclear how the level of integrity was assessed. Moreover, it is suggested that in multifaceted and time-consuming treatments, adherence ratings are expected to be lower than in uncomplicated interventions (Perepletkhikova and Kazdin, 2005). Protocol adherence can be first computed by dividing the number of observed required treatment elements (essential-and-unique, and essential-but-not-unique) by the maximum possible number of these elements (Gresham et al., 1993). Higher scores therefore indicate higher frequencies of essential treatment elements, and thus higher protocol adherence. The second way to determine protocol adherence

is to calculate the percentage of adequate treatment categorizations. Optionally, one can include the unique-but-not-essential and compatible items into the assessment of protocol adherence, but since these are not required, it is difficult to determine to what degree these need to be addressed during treatment.

In our study, in the absence of guidelines, we judiciously determined the cut-off points for sufficient protocol adherence. First, we took into account the fact that both treatments were complex, since these comprised multiple intervention techniques, and were carried out by multiple multidisciplinary therapist teams in several treatment centres. Furthermore, we considered that under some circumstances essential treatment components could be less important (e.g. when behavioural experiments are omitted when catastrophic cognitions cannot be identified), or even irrelevant (e.g. when a session is dedicated to an important event the patient goes through). So, sufficient protocol adherence was defined as at least 70% of essential treatment elements having actually occurred. Additionally, more than 90% of the sessions had to be classified correctly. The occurrences observed of the essential specific treatment elements are displayed in Table 2 to 4. The mean protocol adherence scores, reflected in the mean proportion of essential treatment elements over all evaluated treatment sessions for the phases of EXP as well as GA, are displayed in Table 5.

**Table 5**

Observed protocol adherence (mean proportion of essential treatment elements and correct treatment classification) and contamination (mean proportion of prohibited treatment elements) for the preparatory, educational, and treatment phase of EXP and GA.

	Mean proportion (%) of essential treatment elements (SD)		Correct treatment classification (%)		Mean proportion (%) of prohibited treatment elements (SD)	
	EXP	GA	EXP	GA	EXP	GA
Preparatory phase	86.25 (9.83)	90.83 (11.08)	100.00	100.00	3.00 (4.83)	1.25 (3.95)
Educational phase	65.34 (11.18)	67.19 (19.55)	87.50	100.00	4.17 (11.79)	4.17 (7.72)
Treatment phase	67.50 (19.13)	58.33 (15.08)	91.70	100.00	9.03 (9.87)	2.08 (4.87)
Total	73.17 (16.97)	71.53 (20.55)	93.30	100.00	5.72 (9.24)	2.36 (5.44)

Note that no difference was found between treatment conditions in protocol adherence ( $F = .05$ ,  $p = .82$  for the specific treatment elements;  $F = 2.07$ ,  $p = .15$  for treatment categorization) or treatment contamination (Chi-square = 3.07,  $p = .08$ )

ANOVA, with protocol adherence as the dependent variable and treatment condition and phase as fixed factors ( $\alpha = .05$ ), demonstrated that the protocol adherence scores with regard to the treatment elements did not differ significantly between the treatment conditions for each of the phases ( $F = 1.30$ ,  $p = .28$ ), indicating that further analyses could be performed with the mean protocol adherence scores independent of phase. Furthermore, no difference was found in average protocol adherence scores between both treatment conditions ( $F = .05$ ,  $p = .82$  for the specific treatment elements;  $F = 2.07$ ,  $p = .15$  for treatment categorization). It was found that in general 72% (SD =

18.71) of the essential elements occurred during the selected treatment sessions, indicating that the preset criterion of good protocol adherence was met. However, it should be noted that protocol adherence was especially high during the preparatory phase (86-91%), whereas it dropped to 58-68% during the subsequent phases. Additionally, both independent raters classified the recorded sessions to the correct treatment condition in 97% of the cases: each rater incorrectly classified one EXP session as GA. This indicates that both treatments were easily differentiated from each other, which supports protocol adherence in both treatment conditions.

### **Step 6. Calculation of treatment contamination**

An a priori level of the degree in which prohibited treatment elements are allowed during treatment is needed to consider treatment contamination absent (Pereplechikova and Kazdin, 2005). Again, to our knowledge, no guidelines are available for this. Treatment contamination can be computed by dividing the number of observed prohibited treatment elements by the maximum possible number of these elements, with higher scores indicating higher frequencies of prohibited treatment elements, and thus higher degrees of contamination.

In our study, sufficiently low treatment contamination was defined as maximally 10% of prohibited treatment elements being actually addressed during treatment. Since contamination sum scores per session were not normally distributed, with hardly any values larger than 0, both parametric and non-parametric testing were inappropriate. We therefore dichotomized the contamination sum scores by recoding scores above 0 into 1, and subsequently performed separate Chi-square tests between treatment conditions in general or per phase of treatment ( $\alpha = .01$  because of multiple testing). The occurrences observed of the prohibited specific treatment elements are displayed in Table 2 to 4. The mean contamination scores for the phases of EXP as well as GA are displayed in Table 5. Contamination scores did not differ significantly between the treatment conditions for each of the phases ( $\chi^2$ 's  $< 4.44$ , p-values  $> .04$ ), thus further analyses could be performed independent of phase of treatment. Also, contamination scores did not differ between both treatment conditions ( $\chi^2 = 3.07$ ,  $p = .08$ ). It was demonstrated that overall 4% (SD = 7.71) of the prohibited elements occurred during the evaluated treatment sessions, indicating that treatment contamination was sufficiently absent. For each of the treatment phases, the mean contamination score was below the predefined 10%.

## **CONCLUSION AND DISCUSSION**

Based on available recommendations in the literature, we succeeded in designing, as well as feasibly applying, a method to measure protocol adherence and contamination in a randomized controlled trial comparing two behavioural treatments for chronic low back pain. Sufficient agreement between two experts and a group of trained therapists in evaluating the specific items as being essential/essential-and-unique/unique-but-not-essential/compatible/prohibited supported the content validity of the MPAC. The reliability of the application of the MPAC was supported by adequate interrater reliability of two independent raters. From their ratings of a selection of recorded treatment sessions, it was shown that protocol adherence was sufficient and that treatment contamination was almost absent, at least according to our a priori criteria. The sessions from both treatment conditions were also easily distinguishable from

each other. Besides the importance of establishing sufficient protocol adherence and the absence of treatment contamination, another relevant finding was the equal protocol adherence and contamination in both treatments. This result at least ensures that neither of these treatment conditions was favoured by being better conducted as compared to the other. The comparisons between the effectiveness of EXP as compared to GA therefore seem fair.

Whereas this study put a great effort in quantitatively addressing protocol adherence and contamination based on available recommendations in the literature, this procedure should not be interpreted without acknowledging some drawbacks. There are four main limitations that need mentioning. First, it is intricate to predefine criteria for adequate protocol adherence and treatment contamination in the absence of clear guidelines. In our study, it was challenging to predetermine how many essential elements should have occurred for adequate protocol adherence, especially considering that under some circumstances essential treatment components could be less important or even irrelevant (see 'Calculation of protocol adherence'). Furthermore, retrospectively, we may even have overestimated the importance of several items. For example, during the education phase, the low occurrence of items 1, 9, 14, and 17 may be explained by the fact that these have become abundant since the rehabilitation physician previously explained the bio-psycho-social view of chronic pain, which was in fact in accordance with both our treatment protocols. For future designs of MPAC's, we recommend to only select those specific ingredients that are essential irrespective of circumstances. Furthermore, we did not integrate the specific treatment elements that were compatible or unique-but-not-essential in the calculation of protocol adherence or treatment contamination, since it is difficult to determine to what degree these need to be addressed during treatment. We therefore decided that these did not have any additional value, at least for our treatment outcome study that after completion of the trial verified whether protocol adherence was sufficient and equal between both treatment conditions. Forthcoming studies may individually decide whether these treatment elements are relevant for their assessment of protocol adherence.

Second, depending on the aim of the evaluation of protocol adherence and contamination, the manner in which these factors are assessed and evaluated differs. Since it was our aim is to evaluate whether treatment comparisons were fair, in the sense that both treatments were adequately and equally delivered, we averaged protocol adherence and treatment contamination over all evaluated sessions. However, by doing so, we disregarded the variability in protocol adherence and contamination between sessions as well as between treatment elements. Thus, while we concluded that in general protocol adherence and treatment contamination were regarded as adequate, examination of the various phases of the treatments shows that protocol adherence was especially high during the preparatory phase, whereas it dropped below the predefined criterion during the subsequent phases. Furthermore, although the therapists adhered to some treatment elements in almost all of the cases, this was insufficient in case of other treatment elements. On the other hand, when the aim is to enhance protocol adherence and minimize treatment contamination, these aspects should not only be assessed repeatedly during the trial, but also detailed information about adherence and contamination in each session and per treatment element will be

more suitable. Also for future users of the therapies, detailed information about treatment elements and different phases may be relevant.

Third, the mere assessment of protocol adherence and treatment contamination disregards the evaluation of treatment competence (Waltz et al., 1993), treatment receipt by the patient, and enactment of the patient upon treatment (Lichstein et al., 1994; Bellg et al., 2004). In fact, it is possible that, despite adequate protocol adherence, treatment competence is compromised, as are the understanding and implementation of the treatment by the patients. To fully understand, and compare, the effects or treatments, these components have to be encouraged and evaluated as well. Nevertheless, completion of the Treatment Fidelity Checklist (Borrelli et al., 2005) revealed that our study applied 17 out of 25 (68%) treatment fidelity strategies, which is higher than the reported average 55% of other studies: we utilized 100% of the treatment fidelity strategies of treatment design, 75% of training providers, 80% of delivery of treatment, but none of those from receipt of treatment and enactment of treatment skills.

Fourth, therapist's awareness of the recording of their treatment sessions may influence their behaviour. It is therefore possible that protocol adherence is optimized and treatment contamination minimized in case of the sessions observed, and that because of this, treatment fidelity ratings are inflated (Peterson et al., 1982; Moncher and Prinz, 1991; Gresham et al., 1993; Perepletchikova and Kazdin, 2005). In fact, in our study this might even be more of an issue, since only 21% of the treatment sessions were recorded despite repeated instructions and reminders that the therapists had to record as many treatment sessions as possible.

Despite these limitations, this study presents several important steps derived from available recommendations in the literature that can be undertaken when developing and applying an MPAC, and comprehensively illustrates these subsequent steps within a randomized controlled trial in the area of behavioural medicine. Presentations of methodological issues and detailed examples of these methods may encourage the assessment of protocol adherence as a partial reflection of treatment fidelity in treatment outcome studies. Researchers preparing forthcoming studies may use these to develop their own MPAC, although these methods obviously have to be modified according to the contents of the treatments of interest, and the aims of the assessment.

## ACKNOWLEDGEMENTS

This study was supported by Grant nr. 1436.0002 of the Netherlands Organisation for Health Research and Development (ZonMw). We are grateful to the many therapists and support staff from the treatment centres of Maastricht, Hoensbroek, Sittard and Eindhoven, who performed the treatments and recorded the treatment sessions, to the patients willing to participate and giving consent for treatment recordings, to Sita van Riet for her assistance in the data collection, and to Sarah Holla and Dominique Gielen for their evaluations of protocol adherence and contamination.

# Part V

## Summary & General discussion





## **Chapter 9**

### **Summary & General discussion**



Low back pain (LBP) is a very common health care problem. Fortunately, most people recover from acute LBP within few weeks, and only a small minority (about 10%) of these people develop enduring complaints. Generally, chronic low back pain (CLBP) cannot be adequately explained by biomedical pathology. A better account for these chronic complaints might be provided by a biopsychosocial model that integrates psychological and social factors with biomedical ones. One prominent example of such a model is the fear avoidance model that proposes that pain-related fear might be the central mechanism in the development and continuation of complaints in some chronic low back pain (CLBP) patients. Support for this model has already been established several years ago (Vlaeyen and Linton, 2000), and it has even been stated that ‘pain-related fear is more disabling than pain itself’ (Crombez et al., 1999b).

The objectives of this dissertation are to elaborate on, and test, the role of pain-related fear in CLBP patients, to evaluate various assessment methods of pain-related fear, and to test implications of pain-related fear for the clinical management of CLBP. To contribute to this purpose, this dissertation is divided into these three sections that jointly present seven studies. In this chapter, first a summary is presented of the specific study aims and the main findings that are derived from these studies, for each division of the dissertation consecutively. Moreover, these findings will be presented in light of other available research on these subjects. After this, some methodological limitations of these studies are discussed, as are recommendations for forthcoming studies that can be derived from these. Finally, several recommendations for the clinical management of pain-related fear in CLBP patients will be put forward, as well as some remaining recommendations for future research. Finally, in the end of this chapter some tentative conclusions are presented.

## **SUMMARY OF STUDY AIMS and MAIN RESULTS**

### **Part II: Outlining and testing the contribution of pain-related fear in CLBP**

Several years ago, it was already concluded that ‘the evidence gathered in the last decennium favours a cognitive-behavioural model first forwarded by Lethem et al. (1983) and later, in refined version, put forward by Vlaeyen et al. (1995a)’ (Vlaeyen and Linton, 2000, p. 329). In the first part of this dissertation, additional support for the fear avoidance model is searched for.

#### **Is the fear avoidance model in CLBP patients corroborated by current scientific evidence?**

An update of the Vlaeyen & Linton review (2000) addresses various interrelationships as proposed by this model, as well as the contribution of pain-related fear in the initiation of LBP and the development and maintenance of CLBP. Review of these recent research studies in chapter 2 adds to the already obtained support for the interrelationships of the fear avoidance model of chronic musculoskeletal pain by the following findings: (1) Personality factors, such as trait anxiety or negative affectivity (or neuroticism), may predispose a person to catastrophically (mis)interpret pain and become fearful of pain; (2) CLBP patients who catastrophize about their pain may be

more likely to suffer from more intense pain, increased disability, and higher levels of pain-related fear; (3) Especially CLBP patients with heightened levels of pain-related fear and pain catastrophizing may demonstrate attentional disruption due to pain, resulting from difficulties in disengaging attention from pain and pain-related stimuli; (4) Highly fearful CLBP patients may avoid activities expected to be painful or harmful, and may therefore demonstrate submaximal performance on various behavioural tasks, and report higher levels of disability, than low fearful CLBP patients; (5) In CLBP patients, neither lower physical activity levels nor the physical consequences of long-term activity avoidance are unambiguously confirmed; (6) The degree to which a person feels disabled due to acute or chronic LBP may be significantly determined by the severity of the pain experienced; (7) Pain-related fear may be associated with the inception of LBP in currently pain-free people, with the development of long-lasting disability out of an acute episode, and with the maintenance of chronic pain and disability in CLBP patients.

These findings suggest that the fear avoidance model may adequately describe the process through which a selective group of people fail to recover from an acute LBP episode and develop enduring complaints. However, several precarious points require mentioning. *First*, the fact that CLBP patients subjectively report to experience functional disability and activity limitations is, to the best of our knowledge, not consistently corroborated by objective measurements of decreased physical activity levels (Verbunt et al., 2001; Spenkelink et al., 2002; Bousema et al., 2007; van den Berg-Emons et al., 2007) or deteriorated physical condition (Wittink et al., 2000; Wittink et al., 2002; Brox et al., 2005; Smeets et al., 2006c). *Second*, whereas most studies examined one or several interrelationships of the fear avoidance model, none investigated the validity of the complete model as a whole, nor are its causal mechanisms established. Even the two studies to date, examining the interrelationships of various elements of the fear-avoidance model simultaneously using structural equation modelling, failed to include pain vigilance and avoidance behaviour (Cook et al., 2006), and avoidance and disability (Goubert et al., 2004c). *Third*, this model presents a simplified mechanism that predicts that once patients respond to pain with catastrophic misinterpretations, they will almost inevitably become mired in a vicious circle leading to chronic pain and disability. However, from a clinical viewpoint, it is difficult to fit patients into this model, since the course of LBP and disability is often recurrent, fluctuating, and erratic. *Fourth*, this model does not explain why some patients persevere in their activities despite pain or fear, the so-called ‘over users’ (Vlaeyen and Morley, 2004).

Future research focusing on the examination of the complete model and its causal mechanisms is therefore warranted to test the model’s validity in explaining the development and maintenance of CLBP in some patients. Until then, the fear avoidance model at least holds promise in light of the many studies supporting its propositions, although its limitations should also be kept in mind.

### **Is fear of movement/(re)injury a mediator between pain catastrophizing and functional disability?**

Another way to seek additional support for the fear avoidance model is to actually test an important proposition of this model, that is whether fear of movement/(re)injury mediates the relationship between pain catastrophizing and functional disability. Since

studies usually include clinical patients, who are likely to differ from the average back pain patient in duration and severity of complaints, the study in chapter 3 incorporates people from the general population who experience (C)LBP. By this means, the external validity of the findings should be enhanced.

In this study, it was found that in people suffering from (C)LBP in the general population, pain catastrophizing at baseline is moderately related to fear of movement/(re)injury six months later, above and beyond the contribution of other relevant co-variables such as the level of fear of movement/(re)injury already present at baseline. This finding corroborates the proposition of the fear avoidance model that people who interpret their pain catastrophically are likely to experience a subsequent increase in fear of movement/(re)injury. Furthermore, it was shown that people with heightened levels of fear of movement/(re)injury or pain severity also report higher degrees of functional disability as compared to patients with lower levels of pain-related fear. However, since pain catastrophizing was not found to be related to functional disability six months later, no evidence was found that the relationship between pain catastrophizing and functional disability is mediated by fear of movement/(re)injury.

These findings concerning the relationships between pain catastrophizing, fear of movement/(re)injury, and functional disability in people experiencing (C)LBP in the general population therefore provide only partial support for the fear avoidance model of CLBP, and leave some indistinctiveness concerning the actual interrelationships between these variables. While still being puzzled by the latter finding, one reason might be that in the general population with LBP complaints, functional disability is mainly determined by the level of experienced pain rather than by pain catastrophizing (Boersma and Linton, 2005a; Peters et al., 2005). Another explanation may be that this population presented with relatively low levels of pain catastrophizing, which may have attenuated the relationships between variables due to reduced variability.

### **Part III: Assessment of pain-related fear**

Whereas at present several major self-report measures assessing pain-related fear are available (McNeil and Vowles, 2004), it can be worthwhile to develop new ones that have additional value to these.

#### **Is the TSK-G suitable to measure fear of movement/(re)injury in the general population?**

In view of previous findings that dysfunctional LBP beliefs in pain-free people are associated with subsequent inceptions of LBP, psychometrically sound instruments are welcomed to assess these beliefs in currently pain-free people. We therefore examined in chapter 4 whether a self-report questionnaire for fear of movement/(re)injury (the Tampa Scale for Kinesiophobia: TSK), that already appeared suitable for CLBP patients, also had adequate psychometric properties when it was adapted for application for pain-free people as well (TSK-G). People from the general population, either who currently experienced LBP complaints or who had not experienced

complaints for over a year, completed the TSK-G in addition to several other health-related questionnaires.

Results described in chapter 4 show that the TSK-G has good psychometric properties both in people with and without LBP complaints. It measures a single factor, and has adequate internal consistency. Validity of the TSK-G is supported by the finding that a higher degree of fear of movement/(re)injury, as measured with this questionnaire, was significantly, though only moderately, related to a higher degree of pain catastrophizing and poorer health status.

The fact that the TSK-G measures a single factor even in people with LBP is at variance with the two-factor structure generally found for the original TSK, reflecting the beliefs that serious medical problems underlie the pain, and that activities would aggravate these problems even further (Clark et al., 1996; Geisser et al., 2000; Goubert et al., 2004d; Roelofs et al., 2004a; Cook et al., 2006; Roelofs et al., 2007). Since we did not verify the fit of the two-factor model in this study, we cannot rule out the possibility that this factor structure would have provided an adequate fit as well.

But what exactly does the TSK-G measure? By rephrasing the TSK items that reflect beliefs when currently experiencing LBP, the TSK-G items represent these same beliefs but then with regard to back pain in general or when the person would have suffered from back pain. Typical expressions of the TSK-G are for example 'if I had back pain...', or 'back pain means...'. For people currently experiencing LBP, the TSK-G probably mirrors the TSK, since the hypothetical LBP is reality for them. For people who do not experience back pain, the TSK-G may reflect their general beliefs that LBP results from injury in the back and that physical activities exacerbate this damage. These beliefs might be the result of possible preceding back pain episodes, or derived from general beliefs in the population (Goubert et al., 2004a).

The TSK-G therefore seems to be a useful and psychometrically sound measure to determine dysfunctional beliefs about LBP in people who at present are pain-free. For acute and chronic LBP patients, the TSK remains to be preferred over the adapted TSK-G. This measure is therefore suitable for inclusion in general population studies examining the contribution of these harm beliefs in the development of future LBP episodes.

### **Is the PHODA-SeV suitable to assess perceived harmfulness of activities in CLBP patients?**

Since the original 100-item Photograph Series of Daily Activities (PHODA: Kugler et al., 1999) is time-consuming and not completely standardized, a shortened electronic version of the PHODA (the PHODA-SeV) was developed as an assessment tool. Advantages of the PHODA-SeV over the original PHODA are its automatic data storage, and the fact that its administration is of shorter duration and more standardized. The study described in chapter 5 investigated whether this novel PHODA-SeV adequately assesses perceived harmfulness of activities in CLBP patients.

Results show that the PHODA-SeV is a suitable instrument to assess the perceived harmfulness of daily activities in CLBP patients. This measure demonstrates adequate internal consistency, test-retest reliability, and stability, with discrepancies between two measurements exceeding 20 points suggesting true change. Its construct validity is supported by the finding that higher degrees of perceived harmfulness of

daily activities are associated with increased fear of movement/(re)injury, pain catastrophizing, functional disability, and pain severity. Validity is further corroborated by the finding that patients who have received exposure in vivo (EXP), that aimed to systematically reduce the perceived harmfulness of activities, show significantly lower PHODA-SeV scores after treatment than patients receiving graded activity (GA) that did not address these assumptions. Although these first results with the PHODA-SeV seem convincing, its psychometric properties require further corroboration by forthcoming studies using larger samples, with various (chronic) pain diagnosis, and more objective behavioural measures to test its predictive validity.

Although this study examined the properties of the PHODA-SeV being applied as an assessment tool, there is no reason to assume that the adequate properties found do not apply when this measure is employed for clinical purposes. The PHODA-SeV might provide more detailed, and useful, information than other questionnaires, such as the TSK. An advantage of this assessment tool is that it can be easily adapted for other musculoskeletal or chronic pain syndromes, by replacing the photographs with others that are relevant for that specific syndrome. Moreover, the specific value of the PHODA-SeV lies in the straightforward construct that it aims to assess, that is the degree that CLBP patients consider various activities to be harmful to their back. Derived from the placement of the photographs, but also from the interaction with the patient in placing these along the harmfulness thermometer, the clinician or test leader is provided with valuable information about the beliefs of the patient with regard to the harmfulness of a variety of activities. For the patient, the method is straightforward and understandable, and much more easy to complete than a series of complex items of certain questionnaires.

Because of the PHODA-SeV being viewed as an idiosyncratic measure, its extremely high internal consistency was unexpected. Although such a high internal consistency may be welcomed, it can also reflect the limitation of the PHODA-SeV in that each item actually reflects the same question. The scores on the PHODA-SeV are therefore not reflective of more complex beliefs, such as that pain signals injury, or that activities need to be avoided. Moreover, its high internal consistency suggests that the PHODA-SeV as a measurement tool for perceived harmfulness of activities may even comprise less items than this shortened version does.

### **Can fear of movement/(re)injury be implicitly measured in CLBP patients?**

Because of some potential drawbacks associated with self-report measures (e.g. their need for conscious introspection and their susceptibility to self-presentational strategies), we aimed to examine whether fear of movement/(re)injury can be implicitly measured. Implicit measures may have additional value because of assessing beliefs in an automatic manner, thus without the need for conscious introspection by the participant. In the study described in chapter 6, CLBP patients as well as healthy controls therefore completed the Extrinsic Affective Simon Task (EAST) and the Go-no-Go-Association-Task (GNAT), which measure automatic associations between back-stressing movements and the evaluation ‘threatening’.

In this study, no evidence was found that CLBP patients differ in their level of implicitly measured fear of movement/(re)injury from healthy controls. Neither CLBP patients nor healthy controls demonstrate stronger associations between back-stressing movements and ‘threatening’ than between these movements and ‘neutral’.



Furthermore, the EAST and the GNAT are found to be unrelated to each other, and to the self-report measure of fear of movement/(re)injury. However, since the reliability of these implicit measures is shown to be poor, it is not possible to draw any conclusions from these findings.

At the time of designing our study and selecting the implicit tasks, most researchers were enthusiastic about the EAST, since this measure enables the assessment of various associations within the same task instead of in a relative manner as is the case in many other implicit measures (de Houwer, 2003a). Back then, hardly any publication was available about the application of the EAST. At present, however, the EAST has been criticized for its poor reliability (de Houwer, 2003a; Teige et al., 2004; Schmukle and Egloff, 2006; de Houwer and de Bruycker, 2007). Moreover, published research studies utilizing the EAST report ambiguous findings, with some corroborating their hypotheses (Ellwart et al., 2005; Huijding and de Jong, 2005; Huijding and de Jong, 2006; Lefaivre et al., 2006; Thush and Wiers, 2007), but others failing in this (Teige et al., 2004; Schmukle and Egloff, 2006; Salemink et al., 2007; de Houwer and de Bruycker, 2007). For example, in the area of chronic pain, Vancleef et al. (2007) succeeded in demonstrating that individuals with high levels of injury/illness sensitivity presented with stronger threat appraisals of pain-related stimuli on the EAST than those with lower levels of injury/illness sensitivity. One major point of criticism on the EAST is that, since responses should be based on the colour of the words, participants do not need to process the meaning of these words. But when not processed, the target concept (e.g. back stressing movements) cannot influence performance, as a result of which the measure will not work as it is intended to (de Houwer and de Bruycker, 2007).

For the Go-no-Go-Association Test, also little research was available a few years ago. To date, there are only two additional studies that implemented the GNAT, but in contrast to our study, they did find positive results with this measure (Loughnan and Haslam, 2007; Teachman, 2007). For example, in the study of Teachman (2007), the GNAT was not only successful in differentiating between high and low spider fearful participants, but it was also related to self-report measures of spider fear, to stress reports when encountering a spider, and to actual touching of the spider.

So, since others studies were able to differentiate between groups and to predict behaviour from the EAST and the GNAT, why did we fail in this? *First*, it might be that the construct of fear of movement/(re)injury cannot be captured in the association between back-stressing movements and the valence ‘threatening’, especially since the threatening words (fatal, warning, dangerous, terrible, mean, and horrible) were unrelated to the back-stressing movements. The choice of others to reflect fear with fearful adjectives (e.g. anxious, scared, nervous) might have been a better choice in this (Ellwart et al., 2005; Teachman, 2007). *Second*, pictures of back-stressing movements might have been more suitable, since these may be better and more salient representations of the actual movements, and thus would have elicited stronger associations than the words that we used in our study (Huijding and de Jong, 2005; de Houwer and de Bruycker, 2007). *Third*, the fact that our participants suffered from chronic pain may have affected the results with these response time paradigms. CLBP patients often show deviant performance on these tasks as compared to healthy controls (e.g. longer response times and more errors), and it is unknown what effect these deviant responses may have on the comparison of compatible and incompatible

trials. *Finally*, other implicit measures than the EAST or GNAT may be more suitable for the assessment of implicit associations between movements and threat, such as the Approach Avoidance Task, comparing responses of pulling a threatening picture towards oneself or pushing them away from oneself, that has proven its ability in implicitly assessing fear (Rinck and Becker, 2007).

To summarize, with respect to these new developments in assessment methods of pain-related fear we can conclude that the TSK-G is a reliable and valid questionnaire of fear of movement/(re)injury in people with as well as without LBP. It is specifically worthwhile for population studies examining fear avoidance beliefs as risk factors for the development of subsequent LBP complaints in currently pain free people. For the assessment of perceived harmfulness of activities in CLBP patients, the PHODA-SeV demonstrates adequate psychometric properties. This renders the PHODA-SeV as a valuable additional tool to existing questionnaires for pain-related fear. However, the implicit measurement of pain-related fear has yet failed to be effective, especially because of the poor reliability of the implicit measures that were used.

#### **Part IV: Clinical implications of pain-related fear for CLBP management**

##### **Is EXP more effective than operant GA in CLBP patients?**

Because of the substantial contribution of pain-related fear in the maintenance of CLBP, it may be an important target for intervention. EXP is therefore proposed as a successful treatment in CLBP patients, and especially in those with high levels of pain-related fear. It was examined whether EXP that systematically aims at diminishing pain catastrophizing cognitions and perceived harmfulness of activities is a more effective treatment than operant GA that does not focus upon pain-related fear.

We were able to show that EXP is an effective treatment, though not more effective than GA, in moderately to highly fearful CLBP patients (chapter 7). It was found that EXP is equally effective as GA in improving functional disability and main complaints, although the group difference directly and six months post-treatment almost reached statistical significance favouring EXP. Both treatments are also equally effective in improving pain intensity and daily activity levels directly and six months post-treatment. The superiority of EXP as compared to GA in altering pain catastrophizing and perceived harmfulness of activities is, however, clearly established.

Whereas our results were not as convincing as the single case studies showing a discernible superiority of EXP over GA (Vlaeyen et al., 2001; Linton et al., 2002a; Vlaeyen et al., 2002a; Vlaeyen et al., 2002b; Boersma et al., 2004; de Jong et al., 2005b), they corroborated those of another randomized controlled trial that also found a trend difference in disability between conditions favouring EXP, despite the finding that EXP was significantly more successful than GA in reducing pain-related fear (Woods and Asmundson, 2007). Several reasons may be considered for the insignificant difference between treatment conditions in our study. First of all, EXP may not have been performed in an optimal form. Although we recognized the importance of the prevention of safety behaviours, which was also touched upon in the

treatment manual (e.g. ‘Safety behaviours are allowed during the first EXP sessions to motivate the patient and to promote successful achievements, but these have to be gradually withdrawn thereafter’ and ‘Beware of safety behaviours and do not miss these’), it may be that these safety behaviours did not gain enough credit during actual EXP delivery. In fact, during the assessment of protocol adherence and treatment contamination as described in chapter 8, we did not consider whether the prevention of safety behaviours was an explicit part of EXP, and not of GA. Due to this, it is impossible to verify whether the use of safety behaviours was successfully prevented during EXP. It may thus be that EXP could have been carried out in a more optimized form, that is including explicit response prevention of safety behaviours in combination with EXP to fear-eliciting activities (Salkovskis, 1985; Rowa et al., 2007).

In addition, the specific treatment skills may have been unevenly distributed across treatment conditions, with EXP in a disadvantaged position as compared to GA. While GA was delivered by physical or occupational therapists who were accustomed to delivering this treatment, EXP was carried out by mini-teams of therapists that were only recently trained for that purpose. Moreover, some of these mini-teams did not gain much practice with EXP treatment. It may therefore be that specific treatment competence, which is an important factor for treatment delivery as well as receipt (Waltz et al., 1993; Lichstein et al., 1994; Bellg et al., 2004; Borrelli et al., 2005; Dobson and Singer, 2005), was inferior in EXP than in GA, possibly resulting in a less favourable outcome for the former than the latter treatment condition.

In addition to possible suboptimal delivery of EXP, the fact that GA may have comprised several favourable aspects may also account for the insignificant difference between treatment conditions. First and foremost, it may be that our control treatment resembled EXP. Before start of treatment, GA patients also received unambiguous assurance by the therapist mini-team and rehabilitation physician that they were able to perform all daily activities, even considering their pain, and when applicable their medical condition. Due to this assurance, as well as medical supervision during therapy by the rehabilitation physician, the threshold to select fear-eliciting activities as treatment goals may have been lowered for these patients. Also, the elaborate explanation of the treatment rationale to the patients, which was also an explicit part of GA, may have been effective in itself in reducing the patient’s levels of pain catastrophizing and pain-related fear (de Jong et al., 2005b). Subsequently, these patients were learned to gradually increase the frequency or intensity of the activities selected, which may bear very close resemblance to the gradual approach of EXP. By performing these activities, patients may have inevitably experienced that they were capable of performing these without the expected adverse consequences occurring. Due to this, patients’ catastrophic cognitions may have been rectified (Burns et al., 2003; Spinhoven et al., 2004; Smeets et al., 2006d), and pain-related fear reduced (Mannion et al., 2001; Klaber Moffett et al., 2004; Woby et al., 2004b). Thus, it might be that GA resembled EXP in the sense that it also exposed patients to fear-eliciting activities, though much less systematically or intentionally as in case of EXP.

Furthermore, small differences between treatments could have been expected in advance because of GA being a bonafide treatment in itself. Meta-analyses comparing (cognitive)-behavioural treatments with active control interventions showed small effect sizes of the difference between interventions (Morley et al., 1999; Hoffman et

al., 2007). Moreover, Wampold (2000) concludes from his meta-analysis on psychotherapy outcome studies that specific treatment ingredients only have marginal contributions to treatment outcome whereas the largest part is due to these “common” treatment factors. So, it does not as much matter which specific treatment is carried out, but rather whether it is competently delivered, and whether both the therapist and the patient believe in the effectiveness of the treatment. When viewing the content of EXP and GA more closely, as is done in chapter 8, it appears that there are many common characteristics between both treatments, and that they only differ on a few unique elements. Moreover, the patients in the GA condition considered their treatment as credible, and had as high expectancies, as the EXP patients.

In sum, both the possibility that EXP was not optimally performed due to insufficient emphasis on response prevention of safety behaviours and poorer therapeutic competence than in GA, as well as GA being a bonafide treatment that may have also comprised valuable EXP components, may have resulted in a smaller contrast between both treatments.

### **Are EXP and GA clinically worthwhile treatments?**

For both treatment conditions, we determined the percentage of patients reporting improvements in main complaints (Patient Specific Complaints) and functional disability (Roland Disability Questionnaire) beyond those that were considered as clinically relevant (chapter 7). From this, it appeared that approximately half of patients reported clinically relevant changes in these outcomes irrespective of treatment condition. Furthermore, we found larger changes with respect to the Roland Disability Questionnaire as compared to another randomized controlled trial (RCT) with comparable patients and settings that concluded that their active treatments were more effective than a waiting list control group (Smeets et al., 2006b). The conclusion therefore seems justified that both EXP and GA appear capable of inducing significant, and clinically relevant, changes in CLBP patients.

### **Is EXP more effective than GA especially in highly fearful CLBP patients?**

It seems logical to assume that EXP is particularly effective when applied in fearful situations, whereas it is not as effective when someone is hardly afraid of activities. So, it was hypothesized that the level of pain-related fear moderates the effectiveness of EXP as compared to GA, in that EXP might be more effective especially in highly fearful patients, whereas smaller differences between treatment conditions might be expected for patients with lower levels of fear. However, since we failed to find any interaction between treatment and perceived harmfulness of activities (chapter 7), no evidence was found for pain-related fear being a moderator of treatment effectiveness. Thus, with the value of EXP as compared to GA seeming independent from the patients' level of pain-related fear, there is no indication that EXP is superior to GA in the subgroup of highly fearful patients. In this respect, the only relevant finding was that in five EXP patients treatment was prematurely terminated by the therapists, because according to their clinical judgment catastrophizing cognitions and pain-related fear were completely absent. From this, it may be concluded that EXP does not seem feasible when pain-related fear is entirely lacking, at least from a subjective clinical viewpoint, since in that case it does not have any starting points for intervention.

So why did we fail to find the anticipated interaction between pain-related fear and treatment effectiveness? The main possible reason for this lies again in the possibility that GA comprised some form of EXP as well. Even for the highly fearful subjects, the explicit assurance that all activities were safe, as well as the gradual increase of activities, may have optimized conditions for these patients to carefully explore their capabilities of performing activities they formerly feared. Thus, derived from our findings it might be concluded that also for highly fearful CLBP patients GA may be as much of a choice as EXP.

### **What are the unique working mechanisms of EXP as compared to GA?**

In addition to comparing the effectiveness of EXP with GA, we also examined whether the effect of EXP as compared to GA is mediated by the working mechanisms that are presumed to uniquely underlie EXP. In chapter 7, we found empirical support for the idea that the effect of EXP on functional disability and main complaints as compared to GA is mediated by decreases in pain catastrophizing and perceived harmfulness of activities. These findings are in line with those of Turner et al., who demonstrated that in chronic pain patients the effect of cognitive behavioural treatment as compared to an education/attention control condition on several pain disability outcomes one year after treatment was mediated by decreases in pain catastrophizing between pre-treatment and post-treatment (Turner et al., 2007).

However, in the absence of a waiting list control group, the current design did not allow for testing the specific mediational mechanisms of each treatment separately. It is therefore not verifiable, but possible, that the effectiveness of GA also partially lies in its ability to alter dysfunctional pain beliefs. Previous studies have indeed demonstrated that the effect of active behavioural treatments is mediated by decreases in pain catastrophizing (Burns et al., 2003; Spinhoven et al., 2004; Smeets et al., 2006d). This indicates that even treatments that do not specifically target at correcting catastrophizing cognitions about pain and activities, but that are directed at improving patients' physical activity levels, are capable of modifying these beliefs.

### **To summarize, does EXP has supplemental value for (fearful) CLBP patients?**

The previously mentioned possible reasons for the fact that EXP appeared equally effective as GA all implicate that the contrast between both treatments was smaller than anticipated. Hence, differential treatment effects were more difficult to detect. In face of an almost significant group difference in favour of EXP together with a probable power problem (see subsection 'power' in the section 'methodological considerations and recommendations') makes it tempting to conclude that EXP would have been found to be more effective than GA in case of sufficient statistical power. Still, even if we had found a statistically significant difference, the difference would have been small (unadjusted effect size around .20), indicating that EXP would only have had a marginal additional value to GA in reducing disability and complaints. The supplementary value of EXP as compared to GA holds, however, when the aim of treatment lies in diminishing pain catastrophizing and perceived harmfulness of activities.

### **Were EXP and GA carried out with sufficient treatment fidelity?**

To be able to make honest and genuine comparisons between effectiveness of treatments, the fidelity of the treatments being compared should be assured. In our RCT comparing EXP with GA in chapter 7, we therefore verified whether both treatments were carried out according to the treatment manuals, that is whether protocol adherence was sufficient and treatment contamination absent, as a partial reflection of treatment fidelity.

Protocol adherence and treatment contamination were successfully assessed with a self-developed instrument (Method for assessing Protocol Adherence and Contamination = MPAC), for which adequate interrater reliability and content validity were established (chapter 8). Equal levels of protocol adherence and treatment contamination were found for EXP and GA. Furthermore, in general, protocol adherence appeared adequate since 72% of the essential treatment elements occurred during treatment, and treatment contamination absent since 4% of the proscribed elements took place. These findings provide assurance that outcome comparisons between these treatments are not affected by inadequate or imbalanced treatment delivery. Furthermore, our detailed illustration of methodological issues encountered in designing and applying an MPAC may encourage the assessment of protocol adherence as a partial reflection of treatment fidelity in forthcoming treatment outcome studies.

As our study is one of the few that describes the assessment of protocol adherence and treatment contamination in a detailed manner, we are not able to compare our findings with those of others in this respect. In interpreting our results, some limitations of the procedure applied should be acknowledged, and these will be outlined in the following ‘methodological considerations’ section of this chapter.

## **METHODOLOGICAL CONSIDERATIONS and RECOMMENDATIONS**

### **Selection of participants**

In the studies described in chapter 5 to 7, patients were included that were either referred for participation, or who actually participated in the RCT comparing EXP with GA. By selecting participants on the basis of several selection criteria, we might have included a selective sample of CLBP patients that is not representative of CLBP patients in general. However, most of our criteria are reasonable (contra) indications for rehabilitation treatment, and will therefore also apply in general rehabilitation practice. The only selection criterion that may have restricted our sample is the fact that the patients had to report at least moderate fear of movement/(re)injury, as was indicated by a score of 34 or higher on the TSK. As this corresponds to the second decile or higher, based on norms of a Dutch CLBP sample (Goubert et al., 2000), we only excluded those patients scoring in the lowest 10%. Since the scores of the patients included varied between the screening procedure and the start of the intervention, examination of their pre-treatment TSK scores showed that these covered, and resembled, the distribution of CLBP patients in general (Goubert et al., 2000). We therefore believe that the selection criteria that were applied in our study did not restrict the sample included.

Most participants of the RCT were already referred to rehabilitation care. But because of recruitment problems, we also searched for participants by advertisements and therefore included some patients who would not have been eligible for rehabilitation care otherwise. It is possible that this group of patients differed from the patients that did actively seek health care, but the small number did not allow for statistical testing. However, a study comparing recruitment of patients with chronic shoulder complaints either by an advertisement or by general practitioners showed no influence of enrolment strategy on the effectiveness outcomes in an RCT (Geraets et al., 2006)

For the study in chapter 3, people suffering from LBP in the general population were enrolled by their completion of questionnaires over a time-interval of six months. Because of a potential selective non-response it can be questioned whether the participants in this study were representative for the general population. Furthermore, the included sample might not be representative for the average LBP sufferer in the general population, since for only less than half of the people sufficient data was available for inclusion in the analyses. Moreover, it was found that the included participants differed significantly from those excluded, by reporting more baseline pain catastrophizing, higher pain intensity scores on both measurements, and more disability at follow-up. Finally, the study sample probably differed in the severity and duration of their complaints from clinical LBP patients who actively seek health care. Therefore, it might be that the conclusions from this specific population cannot be generalized to the excluded group, nor that they can be generalized to the non-responders.

For the study described in chapter 4, people with as well as without LBP from the general population were included in the analyses when they did not have any missing values and when they were aged between 25 and 65, thereby excluding 1424 out of 3664 (39%) respondents. The included group may therefore not have been representative from the group of responders. Furthermore, again, it can be questioned whether those participants who returned the questionnaire were representative of the general population.

## Methods of measurement

### *Accelerometry*

Often a discrepancy is found between self-reported disability and more objective physical measures on which no anomalies become apparent (Mannion et al., 2001; Nielens and Plaghki, 2001; Verbunt et al., 2001). In addition to self-report measures of disability, we aimed to objectively compare the treatment conditions in their ability to increase daily activity levels. The participants in the RCT comparing EXP with GA (chapter 7) therefore wore an RT3 accelerometer, since this device was shown to be a reliable and valid measure of daily activity levels. The analyses with the data collected with these accelerometers appeared rather complex. The main problems that we encountered are outlined below.

First, because of variability between devices, it was recommended to provide each participant with the same accelerometer on repeated measurements (Powell et al., 2003). However, due to defective devices (5 out of 14), and logistical problems attached to too few accelerometers, we were not able to conform to this. In fact, the

same RT3 device was worn on each measurement in 69 % vs. 51 % of the cases of EXP and GA respectively. Because of this, repeated measurements within persons may have become affected by some additional measurement error of the various devices.

Second, only after data collection it was discovered that one of the devices was corrupt. This specific device was of an older type than the others (circuit board number 2VCT0001B in contrast to the newer types with circuit board number 2VCT0001E). There were no indications that the other devices suffered from disrupted recording as well. We therefore decided to exclude the data that were collected with this device.

Third, there were many brief periods during which activity registration was interrupted throughout the day, for example when the patient took a shower. Since the device registered continuously, these interruptions were represented by a series of zero counts. However, we had insufficient knowledge of the reasons and duration of interruption in accelerometer registration either to identify these periods or to be able to impute the missing values with the most accurate estimation of the activity level during that period. In CLBP patients it is furthermore impossible to account for these interruptions on the basis of a period of consecutive zero-counts as was proposed by Masse (2005), since these patients vary largely in resting time. Depending on the reason for this interruption, for example because of resting or swimming, the inclusion of these periods in the calculation of daily activity levels may have either underestimated or overestimated the activity level. Furthermore, days on which patients indicated to have abstained from wearing the device because of illness were not included in the analyses. However, it might be argued that sickness is related to CLBP, since people with more severe complaints might also be sick more often, or because people may have reported illness when in fact they experienced severe LBP that forced them to rest. Exclusion of these days may therefore have resulted in an overestimation of the daily activity levels. Of course, imputation of these missing data with the most feasible estimation of the actual activity level would have been the most decent way in the calculation of daily activity levels, but was far too intricate for the current study (Catellier et al., 2005).

Fourth, only data of relatively few people were included in the analyses, since much data was either missing or excluded from the analyses. The majority of the data was forfeited because people did not adhere to wearing the accelerometer during the whole day. Due to the strict criterion for a measurement day to be considered as a complete day, and the fact that at least five complete days of measurement were required per patient, a large proportion of people failed to fulfil these criteria. In addition, patients treated and measured in rehabilitation center Blixembosch did not participate in this measurement, since we were not able to provide this supplementary test location with their own accelerometers due to financial restrictions. Furthermore, disposal of data collected with the deviant device, and the refusal of some patients to participate in the accelerometry measurement because of considering this to burdensome, decreased the available data even further. Therefore, the analyses were only performed with a subgroup of people, which is in fact 71% at pre-treatment, 55% at post-treatment, and 51% at follow-up 1 assessment. Notwithstanding, no difference was found between the patients included as compared to those excluded from the analyses.



Most importantly, all of these limitations probably have affected both treatment conditions equally, since the same procedure was applied in the analyses irrespective of treatment. So, comparisons between treatment conditions do not seem threatened by the drawbacks of these measures. Moreover, the objective assessment of actual activity levels is that valuable that these restrictions are worth to be accepted.

When including accelerometry in forthcoming studies, it merits the effort to optimize adherence of participants in wearing the activity monitor throughout the day, and for at least five days (for example by explaining the relevance of this to them, by reminders by phone, or by providing an incentive when the participant obeys the instructions). Furthermore, collecting accurate and detailed information about the circumstances during the interruption in activity registration makes imputation of these missing data more achievable (Catellier et al., 2005).

### *Modified version of Credibility-Expectancy Questionnaire*

We made our own translation and modification of the Credibility-Expectancy Questionnaire of Devilly and Borkovec (CEQ: Devilly and Borkovec, 2000), which differed from the translation of Smeets et al. (Smeets et al., In press). We therefore called our version the CEQ-m (see Appendix A in chapter 7 for the Dutch version). Derived from a thorough examination of the original CEQ, in the CEQ-m we considered credibility as the patient's beliefs regarding the effectiveness of the treatment for patients in general, whereas the expectancy was considered to reflect the patient's expectations of treatment for him or herself. Another adaptation was that credibility and expectancy were assessed with regard to the three important domains that the treatments were expected to influence, that is functional disability, daily activity levels, and pain intensity, whereas in the other versions of the CEQ only one domain was assessed ('trauma symptoms' in the original version of Devilly and Borkovec, and 'limitations' in the translation of Smeets). In contrast to these other versions, we could not replicate the two-factor structure comprising credibility and expectancy, but instead found a single factor structure with adequate internal consistency. Since we consider the CEQ-m as a distinct questionnaire, the deviant factor structure in light of adequate psychometric properties does not pose a problem for further analyses. It is, however, difficult to determine with the current data which of both constructs is assessed with our CEQ-m. We therefore refer to this construct as 'credibility/expectancy'.

Another note of importance with respect to the CEQ-m is that its administration was organized by the therapists, and not by an independent research assistant. This may have affected the answers of the patients in expressing their opinion regarding the credibility and expectations of the treatment that was just explained to them by the same therapists. For example, the patients may have been more likely to respond in a socially desirable manner. Notwithstanding, these influences were equal in both treatment conditions.

To our knowledge, we were the first to assess the therapist's credibility of each treatment as well by the use of the credibility items of the CEQ-m (see Appendix B in chapter 7 for the Dutch version). The therapists' opinion of each treatment was analyzed after their training in both treatments, and thus before start of the trial. At that moment, the therapists presented with a significantly higher credibility score for

EXP than GA. The higher credibility of EXP than GA may have influenced the delivery of these treatments, in that the therapists might have carried out EXP with more motivation, enthusiasm, and more persuasiveness than GA. However, it can be questioned whether this remained the case during the trial and the actual implementation of both treatments. During the various supervision sessions, the therapists spontaneously expressed their confidence in GA as well, since they observed improvements in their GA patients that went beyond their initial expectations.

For future treatment outcome studies we recommend assessing the patients' as well as the therapists' beliefs regarding the credibility of the treatments being compared, to verify whether both interventions are considered equally credible, or to correct for differences when present. When the purpose of the treatment is to reduce limitations, the Dutch version of Smeets (In press) can be useful, since the original factor structure comprising both expectancy and credibility (Deville and Borkovec, 2000) was replicated in this translation. However, with our study we demonstrated that the CEQ can also be successfully modified when the aim is to assess credibility/expectancy with regard to various domains. However, in case of significant modifications, the psychometric properties should be ensured.

#### *Assessment of protocol adherence and treatment contamination*

With regard to the assessment of protocol adherence and treatment contamination with the MPAC, some limitations should be kept in mind. First of all, when comparing the therapists' categorization of the treatment elements to those of the two experts to establish the MPAC's content validity, the mean agreement per item was found to be 73%. Based on this adequate mean agreement, we decided to regard the categorization of the experts to be sufficiently corroborated by the opinions of the therapists, and therefore took the experts' categorization as the point of reference to determine for each specific element whether it was essential / essential-and-unique / unique-but-not-essential / compatible / prohibited to occur during each treatment. However, with this reasoning, we ignored the variability in agreement between items. On few items, none of the therapists even agreed with the categorization of the experts. Due to this, it might be that the categorization of the experts was not the best reflection of the 'true value' of some of the treatment elements. Alternatively, one can consider the agreement between the therapists and experts *per item*, and exclude or adapt those items on which the agreement was insufficient.

In addition, we computed the mean protocol adherence and treatment contamination over all evaluated sessions. However, this overall evaluation disregards the variability in protocol adherence and contamination between sessions as well as between treatment elements. Thus, in general, protocol adherence and treatment contamination might have been regarded as adequate, but in some individual sessions the therapists might have adhered better to the treatment manual than in other sessions. This is also visible in the differences in protocol adherence and treatment contamination during the various phases of the treatments (see table 5 of chapter 8): protocol adherence was especially high during the preparatory phase (86-91%), whereas it dropped to 58-68% during the subsequent phases. Furthermore, although the therapists adhered to some treatment elements in almost all of the cases, this was

insufficient in case of other treatment elements. For example, for EXP it was found that in 50% of the cases the treatment element ‘the therapists pay a lot of attention to the patient’s pain behaviour’ occurred during treatment delivery, while in fact this was prohibited. Since it was our aim to verify whether treatment comparisons were fair after completion of the trial, the mere examination of average scores over all evaluated sessions can be justified. However, for other applications this will neglect important aspects of treatment delivery. It may have been more suitable if we had made the effort to repeatedly monitor and evaluate protocol adherence and treatment contamination already during the execution of our trial, and to intervene where appropriate to optimize these factors (Cook and Campbell, 1979; Moncher and Prinz, 1991; Gresham et al., 1993; Bellg et al., 2004; Perepletchikova and Kazdin, 2005).

Although we urge forthcoming studies to verify adequate and equal fidelity of their treatments being compared, it should be acknowledged that the complete concept of treatment fidelity is rather complex to assess, since it comprises the adequacy of treatment delivery, the competence of the treatment agents, the understanding of the treatment by the patient, and the ability of the patient to integrate the newly learned techniques into daily life (Borrelli et al., 2005; Perepletchikova and Kazdin, 2005). For each study, it should be determined which components of treatment fidelity are achievable to assess, by considering the importance of the specific components, and the logistical limitations of that study. Furthermore, forthcoming studies may want to evaluate and intervene early on treatment delivery during their trial, to optimize protocol adherence and minimize contamination with prohibited elements by feedback.

## Evaluation of treatment effectiveness

### *Absence of a waiting list control group*

We did not include a waiting list control group in addition to our two treatment conditions, since the effectiveness of cognitive behavioural treatments is already established (Morley et al., 1999; Guzman et al., 2001; van Tulder et al., 2001; Ostelo et al., 2005; Hoffman et al., 2007). In abstaining from this, the conclusions to be drawn about the effectiveness of each treatment, as well as regarding possible moderators and mediators for the treatments separately, are restricted.

Because we wanted to make some deductions regarding the effectiveness of EXP and GA separately, we calculated the percentage of CLBP patients per treatment conditions that reported improvements on the primary outcomes that can be considered as clinically relevant. However, in view of a lack in criteria that reflect clinically relevant improvements beyond those that are statistically estimated, we should be cautious with drawing conclusions about the success of interventions in inducing clinically relevant changes. It is therefore possible that we concluded that a clinically relevant improvement has taken place in a certain patient, whereas this would not have been confirmed by the opinion of the patient self. Furthermore, we compared the decreases in our outcomes with those reported in a comparable study that established the effectiveness of three active treatments as compared to a waiting list (Smeets et al., 2006b; Smeets et al., 2007). Since we found larger decreases in both treatments conditions on the Roland Disability Questionnaire, we conclude that both

EXP and GA produce significant changes in disability. However, it should be acknowledged that we could only determine that our study was comparable to the one of Smeets et al. based on known parameters, whereas it is feasible that other unidentified variables may have been of influence as well.

More effective treatments are likely to be developed when we gain more insight into which treatment works for whom (moderators), as well as in the working mechanisms of these treatments (mediators) (Kazdin and Nock, 2003; Vlaeyen and Morley, 2005; Turner et al., 2007). Since we were not able to investigate moderating variables separately for each treatment condition in the absence of a waiting list, we restricted our moderator analyses to the examination of whether EXP would especially be more effective than GA in highly fearful patients (e.g. whether the difference between treatment conditions interacted with the level of perceived harmfulness of activities). Also, mediators of effectiveness could not be examined for each specific treatment condition, but only for the treatments as compared to each other. With regard to the mediating variables, we therefore confined our analyses to only those unique mechanisms that were assumed to be responsible for the effectiveness of EXP as compared to GA, that is pain catastrophizing and perceived harmfulness of activities.

Although it can be suggested that the inclusion of a waiting list control group is unethical when the effectiveness of the examined interventions has been clearly established, it does enable future studies to draw several important conclusions that are otherwise not feasible. Not only does it allow for conclusions about the effectiveness of each treatment separately, but it also enables the examination of mediators and moderators of the effectiveness of each treatment. Forthcoming studies ought to invest in the assessment of moderating and mediating variables as much as possible, since these provide the much-needed insight into ‘what treatment works for whom and why’ (Vlaeyen and Morley, 2005). Even though the absence of such a control group confines the scope of the analyses, moderating effects can be examined by studying interactions between presumed moderating variables and the treatments, whereas mediating effects can be studied by examining whether the effectiveness of one treatment as compared to the other is mediated by putative unique mechanisms of that treatment. Future studies should therefore carefully consider the benefits and costs of a waiting-list control condition.

### *Power*

The power calculation that we performed mainly relied on the effect of EXP as compared to GA or a no-treatment period as reported by previous studies using replicated single case experimental designs (Vlaeyen et al., 2001; de Jong et al., 2002; Linton et al., 2002a; Vlaeyen et al., 2002b; Boersma et al., 2004). These studies demonstrated substantial effects of EXP, as they did not find any improvement in the GA condition. To make a more realistic estimation, we contrasted the improvements found with EXP in the single case studies to the improvements reported for GA in a RCT of van den Hout (2002). Still, the hypothesized effect size was large ( $d = .68$ ), so in retrospect it can be argued that this concerned an overestimation of the difference between EXP and GA. Moreover, especially in the view of the inclusion of less participants than required based on the power calculation (85 instead of 110), the power of this study is probably too low. It also needs mentioning that the power

calculation was only performed with respect to the group difference after treatment, and thus without considering whether the power was sufficient to additionally determine mediating and moderating effects.

### Statistical analyses

We utilized various statistical techniques for the different studies. Some of these need further consideration. First, with regard to the study assessing the role of fear of movement(re)injury in the general population (chapter 3), the inclusion of only those participants with less than 10% of missing data can be criticised. Not only did this result in the majority of people being excluded from analyses (244 out of 396, which is 62%), but also were the missing data of those included substituted by the questionnaire specific mean of that person. The same applies to chapter 4 that eliminated all people with missing data from the analyses. More sophisticated imputation methods are available, in addition to methods that do not even require imputation of missing data (Schafer and Graham, 2002; Moerbeek et al., 2003). Furthermore, confirmatory path analysis might have been more suitable than exploratory regression analyses, since we had clear hypotheses regarding the interrelationships of these variables (Streiner, 2005; Edwards and Lambert, 2007).

With respect to the RCT (chapter 7), it can be criticized that we tested the effect of treatment for the post-treatment and follow-up 1 assessment combined, since this may disregard the differential effect on either time point. However, the combining of these effects was justified, since for all of the outcome measures the equality of the group effect at both time points was confirmed. Furthermore, combining this effect resulted in an increase in statistical power, which was welcomed in view of limited sample size. Notwithstanding, whether combining of the treatment effect is prudent is open to subjective statistical preferences.

The mediation analysis in our RCT was limited by examining the mediating mechanisms in a time-dependent manner, meaning that the baseline mediator value served as covariate for the baseline outcome measurement, the post-test mediator value as covariate for the post-test outcome, and the follow-up 1 mediator value as covariate for the follow-up 1 outcome. It would be interesting to repeat these analyses in a cross-lagged manner, to examine whether decreases in the mediators between pre- and post-treatment are predictive of future improvements in outcome, and not vice versa (Burns et al., 2003; Turner et al., 2007). We are currently preparing such analyses with data on pain catastrophizing, pain-related fear, functional disability, and pain intensity, which were collected on a daily basis during treatment.

### Generalization to clinical practice (external validity)

It is important to address the question whether the findings of the RCT (chapter 7) can be generalized to the treatment of CLBP patients in general rehabilitation care (Green and Glasgow, 2006). This pertains to whether we examined the *efficacy* of EXP and GA, that is the impact of the specified treatments by controlling as many sources of variability as possible, or their *effectiveness*, thus whether both treatments are feasible and beneficial in other environments as well (Morley and Williams, 2002). We believe that the findings of our RCT are more reflective of their “effectiveness” than their “efficacy”. The reasons for this opinion are as follows (Tunis et al., 2003). *First*, as we already elaborated on, the participants seem representative for the general CLBP

patients. Of additional importance is that the patients were recruited from as many as nine different practice settings, as well as by an advertisement, resulting in a heterogeneous sample. *Second*, the four participating treatment centres seem representative of others in the Netherlands. *Finally*, the treatments under examination, including the control condition, are relevant for clinical practice and can be easily integrated into usual treatment programs.

### *Selection of patients*

Despite these aspects of the trial being representative for general practice, there are also some differences in the selection of patients that should be noted. Not all patients were already referred for rehabilitation care, since we also recruited some patients by an advertisement. These latter patients would otherwise not have been offered rehabilitation treatment, since these persons were not actively seeking health care. In addition, patients were included on the basis of the sole judgment of the rehabilitation physician that rehabilitation would be suitable, in addition to the patients' scores on several questionnaires and questions verified by the research assistants. However, in regular rehabilitation practice, patients are more extensively observed by multiple disciplines that collectively decide whether rehabilitation treatment will be suitable. It is difficult to establish if, and to what degree, this different way of patient selection may have resulted in a deviant patient sample to whom the rehabilitation treatments were offered, and whether this compromises the external validity of the trial.

### *Delivery of treatment*

In addition to differences in patient selection, some differences with respect to treatment delivery need mentioning. *First*, in usual practice, the treatment offered to the patient is tailored to the individual complaints, problems, and goals of that patient. However, in our study it was randomly determined which one of the two treatments the patient would receive. This may have resulted in a *second* distinction, that is in the attitude of the therapists with regard to the match between the treatment and patient. Whereas in daily practice, the therapists have a say in the plan of treatment, in the trial they may have been 'forced' to offer a treatment to a certain patient which they otherwise never might have chosen. During supervision meetings, the therapists indeed expressed their problems with this. On the other hand, when therapists judged the treatment approach to be completely inappropriate, the treatment was terminated and another one offered. *Third*, in usual care there is room for deviations from the planned treatment when during therapy problems are encountered which require attention. In contrast, in our trial, the content of the assigned treatment was fixed by the treatment manual, as a result of which the management of problems that did not pertain to the treatment protocol was postponed until after this treatment. *Fourth*, whereas work resumption might be a frequently encountered goal in regular rehabilitation care, our treatments did not specifically aim at this. However, work resumption could be integrated in both treatments when wished for by the patient, as long as this was strived for by means of the specific working mechanisms of the treatment. *Finally*, a substantial difference may lie in the total amount of treatment delivered. Most rehabilitation programs are quite extensive and intensive, and usually comprise of patients spending complete days in various kinds of therapy, whereas the treatments in our trial only consisted of 16-26 sessions that were delivered twice a

week in one-hour sessions. Again, it is intricate to determine whether these differences in treatment delivery between our trial and usual practice compromise the external validity of the findings.

### *Implementation of treatments*

In order for a novel intervention to be worthwhile being implemented into usual care, information is required about its cost-effectiveness. For this, not only the effectiveness of the treatment in terms of benefits for the patient is considered (e.g. reduced functional disability, improved quality of life), but also the societal costs associated with delivering that treatment (e.g. therapist occupation, and use of facilities in the treatment center), as well as costs borne by the patient, such as costs of health-related activities (e.g. health care utilization), out of pocket expenses (e.g. professional home care), and production losses due to sick leave. In view of no significant difference in effectiveness between GA and EXP, it is relevant to examine which one of these treatments is the least expensive, since this one should be preferred over the other. We are currently preparing such analyses. When from this EXP would appear to be the most cost-effective treatment that is worth implementing into usual care, we do not foresee many difficulties because of smooth implementation of this treatment into the four participating treatment centres in the RCT. Detailed treatment manuals of EXP (Vlaeyen et al., 2004a; Leeuw et al., 2006) are available in multiple languages to inform upcoming therapists about this treatment. However, additional training and sufficient practice in this protocol is highly recommended.

### *Applicability of EXP in other chronic pain diagnoses*

With EXP being a successful treatment for CLBP patients, is it also a suitable treatment in other chronic pain syndromes? Since pain-related fear is also found to contribute to complaints in other chronic pain diagnoses, such as chronic headache (Norton and Asmundson, 2004; Nash et al., 2006), fibromyalgia (de Gier et al., 2003; Turk et al., 2004), whiplash disorder (Sterling et al., 2003; Nederhand et al., 2004), and chronic fatigue syndrome (Silver et al., 2002; Nijs et al., 2004), it can be argued that it might be. In tailoring EXP for each of these diagnoses, it may be important to identify the current concerns for each specific diagnosis, and to contemplate whether the intervention should be targeting proprioceptive stimuli, such as movements in CLBP patients, or possibly interoceptive stimuli (e.g. fear for the pain experience itself). Until now, preliminary support for the effectiveness of EXP is already found in patients with Chronic Regional Pain Syndrome type I (de Jong et al., 2005a).

## **IMPLICATIONS and REMAINING RECOMMENDATIONS**

### **Clinical implications**

Because of pain-related fear being a significant contributor to complaints during various stages of LBP (see chapter 2), it can be argued that this factor and its impact on daily life functioning should be assessed in every patient, and when present, to be integrated in the treatment approach. EXP might be the treatment of choice in reducing pain-related fear in CLBP patients, though comparisons with other interventions are lacking. EXP can be regarded as a tertiary treatment option, since its effectiveness has

thus far been established in chronic patients. However, pain-related fear can also have some important clinical implications for primary (preventing the development of acute LBP) and secondary (preventing the transition of acute to chronic LBP) interventions, which will be discussed after outlining several important implications about the application of EXP in CLBP.

*Clinical considerations about the application of EXP in CLBP patients*

Several important implications about the application of EXP in CLBP are worth considering. First, EXP treatment as was carried out as described in chapter 7 comprises various treatment components that encompass much more than exposure in vivo to fear-eliciting activities. It also consists of elements of cognitive therapy (Beck, 1970; Beck, 1976), such as behavioural experiments and cognitive challenging of catastrophic cognitions, and response prevention of safety behaviours (Salkovskis, 1985). At this time, it is not clear which of the ingredients of the current EXP treatment contributes to its effectiveness. It is therefore not known whether EXP to fear-eliciting activities alone may be sufficient, or that the combination with cognitive therapeutic techniques or behavioural response prevention may have additional value. The fact that GA was (almost) as effective as EXP suggests that the mere exposure to activities that are perceived as harmful may be sufficient to induce significant and relevant changes in functional disability and main complaints. This is supported by other findings that treatments encouraging an increase in physical activity in the absence of cognitive restructuring techniques appear successful in changing maladaptive pain beliefs (Burns et al., 2003; Spinhoven et al., 2004; Smeets et al., 2006d). For example, it was found that a reduction in perceived disability following active therapy was accounted for by reductions in fear avoidance beliefs (Mannion et al., 2001).

Second, the feature of EXP requiring simultaneously a psychologist and a physical or occupational therapist to be carried out, may pose a problem in some clinical practices. However, based on our own accumulating experience, we advise against performing this treatment by a physical or occupational therapist alone. The execution of this treatment might seem straightforward, but it should be recognized that these therapists lack in knowledge and expertise in identifying and addressing cognitive and behavioural elements of human behaviour. It is rather difficult to identify, grasp, and reveal the specific fears and cognitions of these patients who often have no awareness of the existence and content of their worries themselves (Vlaeyen et al., 2002c). In our opinion, the skills of a psychologist who has expertise in cognitive behavioural therapy is therefore probably not only required, but also indispensable for adequate EXP treatment. The same argument applies to the presence of a physical therapist or occupational therapist, who have specific knowledge about the performance of various daily and functional activities, and who uniquely contribute to designing the activity exercises.

Third, an intriguing question is whether EXP is suitable to be carried out in group format. In a group, patients can learn from observing successful EXP exercises in other patients, which may enhance the disconfirmation of dysfunctional pain beliefs and the extinction of fear. We believe, however, that this treatment might not be suitable for group therapy sessions, because of the need for idiosyncratic matching to each patient. The activities that patients fear can vary extensively, in addition to



divergent worries attached to these activities. A group format will inevitably disregard these essential individual differences. One study compared the effectiveness of EXP when carried out in group format added to regular multidisciplinary chronic pain management, and found no additional value of EXP (Bliokas et al., 2007). The fact that their EXP treatment consisted of merely discussing the EXP practices in the group, whereas the actual exercises had to be performed by the patients at home and thus in the absence of the therapist, might explain their null-findings, since patients had every opportunity to avoid activities or to incorporate safety behaviours. In our opinion, the therapists need to be present to monitor whether the practices are adequately carried out.

Fourth, although we failed to find that EXP was more effective than GA specifically in highly fearful CLBP patients, treatment was prematurely terminated in five out of 42 EXP patients because of the absence of pain-related fear. EXP treatment therefore appears inappropriate for patients in whom pain-related fear does not contribute to their complaints. At the same time, however, it is really difficult to identify these patients. During the interview, fear hierarchy, and educational phase, many patients deny to be afraid of movements or pain. In our experience, it is worthwhile not to discard EXP immediately as a treatment option, but to execute a behavioural test, and expose the patient to an activity that most CLBP patients find threatening (e.g. jumping or lifting heavy objects). When patients perform all kinds of activities when asked to, it leaves one with an inadequate foundation for EXP treatment.

Fifth, it can be suggested that a mixture of EXP and GA might be successful, since both EXP and GA appear valuable treatments to reduce disability and complaints in CLBP patients. In the RCT described in chapter 7, EXP and GA were prescribed by treatment manuals. In this respect, EXP has been criticised to be constrained to the performance of threatening activities, since some activities that are not considered as threatening may better be approached in a time-contingent manner with positive reinforcement (e.g. walking). On the other hand, GA has been criticized because of not allowing for the identification and management of pain-related fear in patients in whom it clearly hampers treatment. Based on an adequate functional analysis of the problem, it might be worthwhile to mix GA and EXP elements, and to vary the emphasis on either approach according to the problems presented by the patient. Of course, whether the combination of these treatments works awaits corroboration.

### *Implications of pain-related fear in secondary prevention*

Pain-related fear might be tackled as early as in the acute stage of a LBP episode to prevent the development of a chronic pain problem. It can be suggested that EXP might also be an effective strategy during this stage. Although these results should be interpreted with caution and need empirical support by future studies, it was shown that EXP in addition to usual care was more effective than usual care alone in highly fearful acute LBP patients, whereas EXP had no additional value for acute patients with lower levels of pain-related fear (Sieben, 2005). Another study demonstrated that a 20-minute minimal intervention strategy in general practice, aimed at highlighting the contribution of psychosocial factors in patients with (sub)acute LBP, was especially effective in patients with high levels of fear avoidance beliefs (Jellema et al., 2006). Furthermore, a fear avoidance based treatment, consisting of education

about fear and avoidance in addition to graded exercise, was shown successful in reducing disability as compared to usual care in acute LBP patients high in fear avoidance beliefs (George et al., 2003). Finally, it was demonstrated that even an educational booklet aimed at changing beliefs and encouraging active behaviour in patients with acute or recurrent LBP was effective in reducing fear avoidance beliefs and perceived functional disability (Burton et al., 1999). Although the findings of these studies need replication by future research, they point to the potential worth of reducing pain-related fear in the prevention of chronic complaints.

### *Implications of pain-related fear in primary prevention*

Because of findings showing that fear avoidance beliefs in the pain-free general population might increase the risk of future inceptions of LBP, educating the people about these complaints may prevent subsequent episodes of LBP. It might be worthwhile to change the general beliefs about LBP into a view that it is a very common complaint that almost everyone encounters at least once during his/her life, which is almost always the result of a harmless strain or sprain and not of serious injury, and that can best be managed with staying active in a justified manner instead of resting. Not only will the person affected by an acute LBP episode respond in a more adaptive manner, but the behaviour and beliefs of this person will also be shared and thus reinforced by those of people in his/her environment.

But how should the general view regarding back pain in the general population be modified? The most efficient way to reach as many people as possible is by multimedia campaigns about the causes and management of LBP, which already have proven their effectiveness in reducing fear avoidance beliefs in the community as well as in physicians (Symonds et al., 1995; Burton et al., 1999; Moore et al., 2000). An interesting study in this respect is the state-wide public health media campaign that took place in the state of Victoria in Australia, which aimed to change the attitudes and beliefs about LBP in the general population. It was demonstrated that after this campaign, the beliefs about LBP had changed positively in the general population, in contrast to New South Wales where no such campaign was run. In addition, the attitudes of health care providers were also changed, as was their management of acute and subacute LBP. For example, health care providers were less likely to order tests or to advice work modifications. Furthermore, a reduction in workers' compensation back claims was noticed (Buchbinder et al., 2001). Although there was some cessation in this effect, these beneficial consequences lasted until several years after the campaign (Buchbinder and Jolley, 2005; Buchbinder and Jolley, 2007). This suggests that such a campaign might be capable for a general and long-lasting shift in the attitude towards LBP and its management, both in people from the general population and health care providers.

### **Remaining recommendations for forthcoming research**

In the 'Methodological considerations' section, several recommendations for forthcoming research have already been addressed. In addition to these, some others need mentioning. First, although we examined the effectiveness of EXP as compared to GA until six months after treatment, the success of these treatments in the longer term (e.g. one year follow-up) are essential as well. Additionally, besides analysing the effectiveness of treatment(s) on relevant outcomes, the costs of the treatments as well

as consequent health care expenses are of utmost importance to be able to conclude whether one treatment is favoured over the other(s) in terms of cost-effectiveness. For example, should EXP found to be ‘cheaper’ than GA, it can be concluded that EXP would be favoured in terms of cost-effectiveness as compared to GA. Finally, much more sophisticated examinations of mechanisms of change during treatment are possible by examining the sequence of changes in multiple variables (e.g. whether EXP results in changes in pain catastrophizing, fear of movement/(re)injury, functional disability, and pain severity respectively) on a day-to-day basis. We are currently preparing these abovementioned analyses.

The fact that we almost found a statistically significant superiority of EXP as compared to GA in diminishing functional disability and main complaints in view of a probable power problem, pleads for replication with a large enough number of participants and adequate power. In view of many logistical problems attached to carrying out a multicenter RCT, e.g. with sufficient patient recruiting, studies using replicated single case designs might be valuable and powerful alternatives (Onghena and Edgington, 2005; Morley, In press). In these designs, the effect of treatment is investigated in a small sample of patients, by testing the improvement as compared to a no-treatment baseline period or a contrasting treatment *within* each patient instead of between patient groups.

Furthermore, effectiveness of rehabilitation treatments should be assessed beyond self-report measures of functional disability and complaints, for example by behavioural performance tasks (Smeets et al., 2006a) or physical activity levels (Verbunt et al., 2003a). Treatment outcome studies also ought to ensure the fidelity of the treatments delivered, to optimize the internal and external validity of their findings (Yeaton and Sechrest, 1981; Moncher and Prinz, 1991).

In addition, since generalization of corrective EXP experiences to movements might be difficult to achieve beyond those practiced during therapy (Crombez et al., 2002b; Goubert et al., 2002; Goubert et al., 2005a), it would be interesting to examine how treatment effectiveness might be optimized by enhancing generalization.

Finally, with many research studies examining various interrelationship of the fear avoidance model, the narrative reviews of Vlaeyen and Linton (2000), and the one presented in chapter 2 can be very informative. However, these might be susceptible to confirmation biases, at least when they are not executed according to the guidelines of a systematic review. Furthermore, the mere enumeration of studies confirming a certain relationship is not informative about the magnitude of that relationship, or about the confidence in the conclusions drawn. Forthcoming studies are therefore not only urged to review the studies about the fear avoidance model in a more systematic and objective manner, but also to perform a meta-analysis on the results of these studies, presenting effect-sizes of the associations predicted by the fear-avoidance model (Cook et al., 1995; Thorn et al., 2007).

## CONCLUSIONS

From the studies that are outlined in this dissertation, several, although tentative, conclusions can be drawn. What becomes clear from the review of current scientific evidence for the fear avoidance model of musculoskeletal pain is that many recent studies corroborated the interrelationships as proposed by this model. Furthermore, pain-related fear may be an important contributor in the development of new LBP

episodes, in the transition from acute to chronic complaints, and in the maintenance of complaints once these are enduring. However, there are also studies that are at variance with the fear avoidance model. One of these was our own study, in which we failed to establish that fear of movement/(re)injury mediates the relationship between pain catastrophizing and functional disability in the open population with LBP complaints, as is a specific proposition of the fear avoidance model. Therefore, despite this model holding promise in explaining why specific patients may develop and maintain CLBP, it urges for further corroboration.

With respect to the assessment of pain-related fear, novel developments in self-report questionnaires appear successful. It was shown that in pain-free people, the TSK-G is a suitable measure to assess fear of movement/(re)injury, denoting the belief that LBP results from injury and that activities might aggravate the damage. Furthermore, the PHODA-SeV was shown to be an adequate measure for the assessment of perceived harmfulness of activities in CLBP patients. In contrast, implicit measurements with the EAST and the GNAT failed to be valuable because of their poor reliability. Pending further developments to ensure their psychometric properties, in the future these measures might become valuable additions to self-report measures in the assessment of pain-specific beliefs.

Although EXP appears especially successful in reducing pain catastrophizing and perceived harmfulness of activities, it is not more effective than GA in improving functional disability and individual main complaints. Also no differences between treatments are found in daily activity levels and pain severity directly and six months after treatment. Both treatments are capable of inducing clinically relevant changes in a substantial proportion of patients. In contrast to expectation, EXP is not found to be a more effective treatment than GA especially in highly fearful patients. Treatment comparisons between EXP and GA seem fair based on the verification of therapeutic adherence to the treatment manuals from independent raters' evaluations of objective representations of treatment sessions. Both EXP and GA are thus effective treatments that may be considered in case of CLBP patients, irrespective of their level of fear. However, when the aim is to correct dysfunctional pain cognitions or to reduce fear of movement/(re)injury, EXP might be preferred over GA.



## **Nederlandse samenvatting**

**(Dutch summary)**



## Deel 1: Algemene introductie

In het eerste deel van dit proefschrift, de algemene introductie in **hoofdstuk 1**, wordt allereerst ingegaan op de problematiek van chronische lage rugpijn en op de rol van pijngerelateerde vrees bij deze klachten. Lage rugpijn (LRP) is een veel voorkomend gezondheidsprobleem. Hoewel acute rugklachten bij de meeste mensen vanzelf binnen enkele weken verminderen en verdwijnen, blijft een kleine groep mensen pijnklachten houden. Wanneer de klachten langer dan drie maanden aanhouden, wordt er gesproken van chronische lage rugpijn (CLRP). Vaak gaan de pijnklachten samen met aanzienlijke beperkingen in het dagelijks functioneren. De hoge socio-economische kosten die gepaard gaan met rugklachten worden voornamelijk veroorzaakt door deze groep mensen met CLRP, niet alleen door hun gebruik van de gezondheidszorg, maar voornamelijk door hun ziekteverzuim en arbeidsongeschiktheid. Chronische rugklachten zijn biomedisch vaak moeilijk te verklaren, omdat er in de meeste gevallen geen lichamelijke afwijkingen gevonden kunnen worden die de pijn en beperkingen afdoende verklaren (zogenaamde *aspecifieke* chronische lage rugpijn). Tegenwoordig lijken deze klachten beter benaderd te worden vanuit een biopsychosociale model, dat ervan uitgaat dat de pijn en beperkingen het gevolg zijn van een complexe wisselwerking tussen biomedische, psychologische en sociale factoren. Een vooraanstaand biopsychosociaal model met betrekking tot CLRP is het vrees-vermijdingsmodel. Volgens dit model wordt de ontwikkeling en instandhouding van CLRP in sommige patiënten verklaard door de aanwezigheid van pijngerelateerde vrees. Wanneer mensen een extreem negatieve betekenis toekennen aan de pijn (catastrofaal denken), kan er vrees voor bepaalde bewegingen ontstaan, omdat deze bewegingen geacht worden pijntoename of letsel te veroorzaken. Deze vrees kan aanleiding geven tot een aantal defensieve reacties, zoals vermijding van of ontsnapping aan bedreigende activiteiten, overmatige aandacht voor bedreigende signalen, of veiligheidsgedragingen tijdens de uitvoering van activiteiten. Hoewel deze defensieve reacties leiden tot een reductie in angst en/of pijn, kunnen deze op de lange termijn averechts gaan werken en de vrees in stand houden. De vermijding van bepaalde activiteiten kan op lange termijn leiden tot beperkingen, een verslechterde lichamelijke conditie en een sombere stemming, die samen met overmatige aandacht voor pijn op hun beurt weer kunnen leiden tot een versterkte pijnervaring.

Naast deze inleiding over de problematiek en prevalentie van CLRP en een korte introductie van het vrees-vermijdingsmodel, wordt er in hoofdstuk 1 aandacht besteed aan verschillende bestaande meetmethoden van pijngerelateerde vrees, waarbij zowel enkele zelfrapportage maten beschreven worden als impliciete meetmethoden die associaties op een indirecte wijze trachten te meten. Ook worden er twee vormen van behandeling van chronische lage rugpijn beschreven, te weten operante graded activity (GA) en een cognitief gedragsmatige exposure in vivo (EXP), die beiden worden toegelicht aan de hand van hun onderliggende theoretische achtergrond (respectievelijk de operante leerprincipes en de klassieke conditioneringprincipes). Hoewel beide behandelingen streven naar een vermindering in functionele beperkingen, gaan ze uit van geheel verschillende werkingsmechanismen en technieken. GA leert patiënten hun activiteitsniveau geleidelijk uit te breiden door tijdcontingent te werken, waarbij gezond gedrag positief wordt bekrachtigd en aandacht niet langer uitgaat naar pijngedrag. EXP streeft ernaar om disfunctionele cognities ten aanzien van de schadelijkheid van activiteiten te corrigeren en



bewegingsvrees te reduceren door patiënten bloot te stellen aan activiteiten die de patiënten in oplopende mate bedreigend vinden voor de rug.

Het doel van dit proefschrift is ten eerste om de theoretische basis voor de rol van pijngerelateerde vrees in (C)LRP uiteen te zetten en te onderzoeken, ten tweede om verschillende meetinstrumenten voor pijngerelateerde vrees te testen, en tot slot om behandelimplicaties te onderzoeken die van pijngerelateerde vrees kunnen worden afgeleid.

## **Deel 2: Uiteenzetten en onderzoeken van de rol van pijngerelateerde vrees bij CLRP**

Het tweede deel van dit proefschrift richt zich op het theoretisch uiteenzetten en onderzoeken van de rol van pijngerelateerde vrees in CLRP patiënten.

Omdat onderzoek naar het vrees-vermijdingsmodel in de jaren na een review van Vlaeyen en Linton in 2000 (Vlaeyen J. W. S. & Linton, S. J. (2000). Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. *Pain*, 85(3), 317-332) sterk is toegenomen, wordt in **hoofdstuk 2** niet alleen ingegaan op nieuwe ontwikkelingen ten aanzien van dit model, maar wordt eveneens een uitgebreid overzicht gepresenteerd van de meest recente evidentie voor de verschillende componenten van dit model. In het algemeen worden de veronderstelde relaties van dit model ondersteund door onderzoek: (1) Mogelijk bestaan er bepaalde persoonlijkheidsfactoren, zoals neuroticisme of negatieve affectiviteit, die de kans vergroten dat iemand angstig reageert op een pijnvaring; (2) CLRP patiënten die catastroferen over hun pijn zijn meer geneigd om intensere pijn, meer beperkingen, en meer pijngerelateerde vrees te ervaren; (3) Voornamelijk CLRP patiënten met verhoogde niveaus van pijngerelateerde vrees en pijn catastroferen kunnen verstoorde aandachtsprocessen vertonen, die resulteren van een vertraagde losmaking van de aandacht van stimuli die als bedreigend worden beschouwd; (4) Hoogangstige CLRP patiënten vermijden mogelijk activiteiten die als pijnlijk of bedreigend worden gezien, hetgeen zichtbaar is in een slechtere prestatie op fysieke taken en meer ervaren beperkingen dan door laagangstige CLRP patiënten; (5) Er is geen eenduidig bewijs dat CLRP patiënten daadwerkelijk lagere activiteitsniveaus (disuse) en een slechtere fysieke conditie (fysieke deconditionering) hebben dan gezonde mensen; (6) Pijnintensiteit lijkt in een belangrijkere mate gerelateerd aan ervaren beperkingen bij (C)LRP dan voorheen werd aangenomen. Hoewel de relaties tussen de componenten van het vrees-vermijdingsmodel door verscheidene onderzoeken ondersteund wordt, roepen de ambigue resultaten ten aanzien van de veronderstelde disuse en fysieke deconditionering nog enige vraagtekens op. Daarnaast ontbreken gegevens die aantonen of de validiteit van het model in het *verklaren* van CLRP klachten onderschreven kan worden.

Aanvullend op de evidentie voor de componenten van het vrees-vermijdingsmodel is in dit hoofdstuk de bijdrage van pijngerelateerde vrees tijdens verschillende fasen van LRP beschouwd. Hieruit komt naar voren dat pijngerelateerde vrees niet alleen een mogelijke instandhoudende factor is van CLRP door zijn samenhang met overmatige aandacht voor pijn en vermijdingsgedrag, maar dat het mogelijk ook bijdraagt aan de ontwikkeling van chronische klachten vanuit een acute episode en aan de ontwikkeling van acute LRP in pijnvrije mensen. Pijngerelateerde vrees lijkt daarom een belangrijk aandachtspunt waar behandelingen op aan zouden

dienen te grijpen. In dit hoofdstuk worden daarom eveneens de effectiviteit van verschillende behandelstrategieën uiteengezet, waaronder ook EXP, een behandeling die speciaal ontwikkeld is vanuit de overeenkomsten tussen pijngerelateerde vrees en angststoornissen. Tot slot worden in dit hoofdstuk enkele interessante onderwerpen aangesneden die tot op heden onvoldoende onderzocht zijn.

In **hoofdstuk 3** van dit proefschrift staat een studie beschreven die een belangrijke veronderstelling van het vrees-vermijdingsmodel onderzoekt, namelijk dat pijn catastroferen resulteert in pijngerelateerde vrees, waardoor op termijn beperkingen ontstaan (m.a.w. dat pijngerelateerde vrees een mediator is tussen catastroferen en beperkingen). Omdat voorgaande studies zich voornamelijk hebben gericht op klinische patiënten, die mogelijk niet representatief zijn voor de gemiddelde patiënt, hebben er aan deze studie mensen met rugklachten uit de algemene populatie deelgenomen. In totaal zijn er 152 mensen met LRP uit de algemene populatie geïnccludeerd in de analyses van deze studie. Deze mensen hebben twee maal een batterij vragenlijsten ingevuld met een tussenperiode van zes maanden. Omdat één van de voorwaarden voor mediatie, namelijk dat er een relatie dient te bestaan tussen pijn catastroferen en beperkingen zes maanden later, niet is bevestigd, kan niet geconcludeerd worden dat pijngerelateerde vrees een mediator is tussen pijn catastroferen en beperkingen. Wel werd er een significante relatie gevonden tussen pijn catastroferen en pijngerelateerde vrees zes maanden later, hetgeen suggereert dat mensen die een catastroferende betekenis toekennen aan hun pijn in de toekomst over het algemeen angstiger blijken te zijn dan mensen die niet catastroferen. Daarnaast werd gevonden dat hogere niveaus van pijngerelateerde vrees en pijn intensiteit geassocieerd zijn met meer functionele beperkingen. Samenvattend zijn er in deze studie enige aanwijzingen gevonden voor de rol van pijn catastroferen, vrees en beperkingen in de algemene populatie met rugklachten, hoewel de relaties tussen deze factoren (nog) niet geheel helder zijn.

### **Deel 3: Meetinstrumenten voor pijngerelateerde vrees**

In het derde deel van dit proefschrift worden verschillende meetinstrumenten van pijngerelateerde vrees onderzocht.

In **hoofdstuk 4** staat een studie beschreven die onderzoekt of de Tampa Scale for Kinesiophobia, die reeds geschikt bevonden is om bewegingsvrees in (C)LRP patiënten te meten, ook geschikt is voor het meten van deze vrees in gezonde pijnvrije mensen. Hiervoor is de Tampa Scale for Kinesiophobia-General Population (TSK-G) gebruikt, die zodanig is aangepast dat mensen zonder rugklachten de vragen kunnen beantwoorden. Voor het bepalen van de factor structuur, interne consistentie en validiteit werden 1029 mensen met LRP in het afgelopen jaar en 1211 mensen zonder deze klachten meegenomen in de analyses, die at random waren geselecteerd uit de Nederlandse algemene populatie. Alle mensen hebben de TSK-G ingevuld als onderdeel van een batterij vragenlijsten. Zowel exploratieve als confirmatieve factoranalyse toonde aan dat de TSK-G één enkele factor meet en dat dit geldt voor zowel mensen met als zonder rugklachten in het afgelopen jaar. Deze enkele factor bestaat uit 12 items die tezamen een goede interne consistentie hebben. Ongeacht de aan- of afwezigheid van rugklachten blijken hogere TSK-G scores geassocieerd met hogere scores op pijn catastroferen en een slechtere gezondheidsstatus, hetgeen ondersteuning geeft voor de validiteit van de TSK-G in het meten van bewegingsvrees

in mensen met en zonder rugklachten. De TSK-G lijkt dus een geschikt meetinstrument dat gebruikt kan worden in populatiestudies naar kwetsbaarheidfactoren voor de ontwikkeling van toekomstige LRP episoden.

Een andere conceptualisatie van bewegingsvrees is de perceptie van een patiënt dat activiteiten schadelijk kunnen zijn voor de rug. Dit kan gemeten kan worden met de Photograph Series of Daily Activities (PHODA), die bestaat uit honderd foto's van activiteiten die een patiënt langs een 'schade-thermometer' dient te ordenen naar gelang de activiteiten verwacht worden schadelijk te zijn voor de rug. De originele PHODA versie is echter zeer uitgebreid en de manuele afname is zeer tijdsintensief. Ook zijn er (nog) geen gegevens bekend over de psychometrische eigenschappen van de PHODA, terwijl deze reeds gebruikt wordt in onderzoek en in de EXP behandeling. De studie beschreven in **hoofdstuk 5** onderzoekt daarom de psychometrische eigenschappen van de elektronische verkorte versie van de PHODA: de PHODA-SeV. De PHODA-SeV is herhaaldelijk afgenomen bij CLRP patiënten die deel hebben genomen aan een gerandomiseerd klinisch onderzoek naar de effectiviteit van EXP versus GA. Uit de resultaten blijkt dat de PHODA-SeV één enkele factor meet die een zeer hoge interne consistentie heeft. Ook blijken de test-hertest betrouwbaarheid en de stabiliteit van de PHODA-SeV goed over een tijdsinterval van twee weken, waarbij discrepanties groter dan 20 punten tussen de twee metingen een werkelijke verandering suggereren. De constructvaliditeit van de PHODA-SeV wordt gesteund door significante relaties tussen dit instrument en andere zelfrapportage maten voor vrees, catastroferen, pijn en beperkingen. Wanneer gecorrigeerd voor gedeelde variantie door middel van multiële regressie blijken hogere scores op de PHODA-SeV uniek voorspeld te worden door hogere scores in pijnintensiteit en bewegingsvrees. De validiteit wordt verder ondersteund door de bevinding dat patiënten die de EXP hebben gevolgd, gericht op de systematische reductie van de waargenomen schadelijkheid van activiteiten, significant lagere PHODA-SeV scores rapporteren na afloop van de behandeling dan de patiënten die de GA hebben ondergaan, welke zich niet richt op deze percepties. Concluderend presenteert deze studie aanwijzingen voor de betrouwbaarheid en validiteit van de PHODA-SeV als een meetinstrument voor de waargenomen schadelijkheid van activiteiten in CLRP patiënten.

Gezien het feit dat het meeste onderzoek gebruik maakt van zelfrapportage maten, die niet alleen gevoelig zijn voor o.a. zelfpresentatie strategieën en sociale wenselijkheid, maar die ook introspectie van de patiënt vereisen, wordt in **hoofdstuk 6** een studie beschreven die bewegingsvrees op een impliciete manier tracht te meten. Impliciete maten zouden mogelijk een waardevolle aanvulling kunnen zijn op zelfrapportage maten, doordat deze associaties op een indirecte wijze meten zonder dat de persoon zich bewust hoeft te zijn van datgene wat gemeten wordt. Op deze manier kan bijvoorbeeld het effect van een behandeling op pijngerelateerde vrees bepaald worden zonder dat de meting wordt beïnvloed door sociale wenselijkheid. In deze studie is onderzocht of CLRP patiënten op impliciet niveau meer bewegingsvrees tonen dan gezonde controles. Daarnaast is nagegaan of twee impliciete maten met elkaar en met zelfrapportage maten samenhangen. Hiervoor werden een groep van 66 CLRP patiënten en 30 gezonde controles geïnccludeerd. Aanvullend op zelfrapportage maten werd bewegingsvrees impliciet gemeten met de Extrinsic Affective Simon Task (EAST) en de Go-No-Go-Association Task (GNAT), die de sterkte van de associatie

tussen rugbelastende activiteiten en de evaluatie ‘dreigend’ en ‘neutraal’ bepaalden. De veronderstelling was dat impliciete bewegingsvrees gereflecteerd zou worden in een sterkere associatie tussen rugbelastende activiteiten en ‘dreigend’ dan met ‘neutraal’. Uit de resultaten bleek dat CLRP patiënten niet verschillen in hun impliciete bewegingsvrees van gezonde controles. Zowel de patiënten als de gezonde controles vertonen geen impliciete bewegingsvrees. Ook blijken beide impliciete taken niet gerelateerd aan elkaar of aan zelfrapportage maten voor bewegingsvrees. Echter, de interpretatie van bovenstaande resultaten blijkt bemoeilijkt te worden door de gebrekkige betrouwbaarheid van deze impliciete maten. We raden daarom aan dat er meer onderzoek dient plaats te vinden naar de psychometrische eigenschappen van deze impliciete maten alvorens deze worden aangepast en gebruikt voor het meten van complexe constructen zoals bewegingsvrees.

#### **Deel 4: Klinische implicaties van pijngerelateerde vrees voor de behandeling van CLRP**

In het vierde deel van dit proefschrift worden de behandelimplicaties van pijngerelateerde vrees voor CLRP patiënten onderzocht.

Vanuit de aanwijzingen dat pijngerelateerde vrees een aanzienlijke rol lijkt te spelen in de instandhouding van CLRP lijken interventies aangewezen die deze vrees reduceren. Op basis van duidelijke overeenkomsten tussen bewegingsvrees en andere angststoornissen, alsmede het feit dat EXP een effectieve behandeling is voor de meeste angststoornissen, is deze behandeling ontwikkeld voor toepassing in CLRP patiënten. Tot nu toe hebben kleinschalige single case experimenten de effectiviteit van EXP aangetoond bij enkele hoogangstige CLRP patiënten. In **hoofdstuk 7** wordt in een gerandomiseerd klinisch onderzoek onderzocht of EXP, dat zich specifiek richt op het systematisch reduceren van de waargenomen schadelijkheid van activiteiten, een geschiktere behandeling is voor CLRP patiënten dan de operante GA. Daarnaast wordt nagegaan of deze behandeling voornamelijk effectiever is dan GA bij hoogangstige CLRP patiënten, terwijl een kleiner verschil tussen interventies wordt verwacht voor laagangstige patiënten. In totaal zijn in deze studie 85 patiënten met CLRP met een revalidatie-indicatie en minstens enige bewegingsvrees geïnccludeerd, waarvan er 42 willekeurig zijn toegewezen aan EXP en 43 aan GA. De behandelingen vonden plaats in vier revalidatiecentra en de patiënten werden gevolgd tot zes maanden na afloop van de behandeling. Intention to treat analyses werden toegepast, hetgeen inhoudt dat alle patiënten werden meegenomen in de analyses ongeacht of dat ze de behandeling en/of metingen hadden voltooid. Uit de resultaten bleek dat, hoewel EXP effectiever is dan GA in het reduceren van pijn catastroferen en de waargenomen schadelijkheid van activiteiten, beide behandelingen even effectief zijn in het verminderen van functionele beperkingen en individuele klachten direct en zes maanden na afloop van de behandeling. Het verschil tussen de behandelcondities op deze meetmomenten bereikte bijna het niveau van statistische significantie ten gunste van EXP. Het aantal mensen dat een klinisch relevante verbetering in functionele beperkingen en individuele klachten rapporteerden was gelijk in beide behandelcondities, hoewel voor functionele beperkingen het verschil wederom bijna significant was ten gunste van EXP. Beide behandelingen blijken daarnaast even effectief in het verbeteren van de pijnintensiteit en dagelijks activiteitsniveau. Tegen de verwachting in bleek niet uit de resultaten dat EXP effectiever is dan GA

voornamelijk in de subgroep van hoogangstige patiënten. Uit deze resultaten blijkt dus dat EXP een minstens even effectieve, of mogelijk zelfs betere, behandeling is dan GA voor CLRP patiënten, ongeacht hun niveau van vrees. In de discussie van de resultaten worden meerdere verklaringen gegeven voor het feit dat deze studie niet heeft kunnen aantonen dat EXP effectiever is dan GA. Naast onvoldoende statistische power door inclusie van te weinig proefpersonen zou een belangrijke reden kunnen zijn dat GA essentiële elementen van EXP bevat, zoals de expliciete geruststelling die aan de patiënt wordt gegeven dat alle activiteiten veilig en onschadelijk zijn in combinatie met een geleidelijke reactivatie, en dus blootstelling, aan mogelijk gevreesde activiteiten. Daarnaast bestaat de mogelijkheid dat de behandelaars vaardiger waren in het uitvoeren van GA dan EXP.

Om eerlijke en evenredige vergelijkingen tussen behandelingen te kunnen maken, dienen behandelstudies zo veel mogelijk de integriteit van de uitvoering van deze behandelingen na te gaan. De afwezigheid van een dergelijke controle compromitteert de interne en externe validiteit van de behandelstudie. Ondanks het belang van deze controle, gebeurt dit echter (nog) zelden en is er een gebrek aan gestandaardiseerde methodieken. In **hoofdstuk 8** wordt daarom een gedetailleerd voorbeeld gepresenteerd van de ontwikkeling en toepassing van een instrument dat de protocoltrouw (m.a.w. in hoeverre belangrijke behandelelementen daadwerkelijk aan bod zijn gekomen tijdens de uitvoering van de behandeling) en behandelcontaminatie (m.a.w. in hoeverre behandelelementen, die niet zijn toegestaan tijdens een behandeling, niet aan bod zijn gekomen tijdens de uitvoering van de behandeling) tracht te meten. Dit instrument wordt de MPAC (Method of assessing Protocol Adherence and Contamination) genoemd en is toegepast in het gerandomiseerd klinisch onderzoek beschreven in hoofdstuk 7. Voor de ontwikkeling van de MPAC hebben twee experts op het gebied van EXP en GA verschillende mogelijke behandelelementen geïdentificeerd voor beide behandelcondities. Vervolgens hebben deze experts onafhankelijk van elkaar voor beide behandelcondities apart deze behandelelementen gecategoriseerd in vijf categorieën (essentieel-en-uniek, essentieel-maar-niet-uniek, uniek-maar-niet-essentieel, compatibel, niet-toegestaan). Ook is ditzelfde gevraagd van de behandelaars die hebben deelgenomen aan de studie en die getraind zijn in beide behandelingen. De inhoudsvaliditeit van de MPAC wordt gesteund door voldoende overeenstemming tussen de experts in het categoriseren van de behandelelementen, hetgeen verder bevestigd wordt door de categorisatie van de getrainde behandelaars. De behandelelementen zijn vervolgens in willekeurige volgorde opgenomen in een scoreformulier van de MPAC. De MPAC bevat derhalve een verzameling van behandelelementen van zowel EXP als GA, die variëren in relevantie en uniciteit en waarvan dient te worden aangegeven of deze al dan niet aan bod zijn gekomen tijdens een behandel sessie. Daarnaast vraagt de MPAC om een inschatting tot welke behandelconditie de behandel sessie thuishoort. Twee onafhankelijke beoordelaars die zijn getraind in het gebruik van de MPAC hebben 30 willekeurig geselecteerde audio-opnamen van behandel sessies beluisterd en gescoord met gebruik van de MPAC. De interbeoordelaars betrouwbaarheid van de MPAC wordt ondersteund door voldoende overeenstemming tussen beide beoordelaars in het scoren van dit instrument. Uit de evaluaties van deze onafhankelijke beoordelaars blijkt dat de protocoltrouw bij zowel EXP als GA voldoende is geweest en dat er nauwelijks contaminatie heeft plaatsgevonden. De vergelijkingen tussen de

behandelcondities in de studie van hoofdstuk 7 lijken dus gerechtvaardigd, hoewel hierbij vermeld dient te worden dat het onbekend is of de resultaten, gebaseerd op de opgenomen sessies, gegeneraliseerd kunnen worden naar de meerderheid van de sessies die niet opgenomen zijn. Deze studie biedt een gedetailleerde illustratie van het ontwikkelen en toepassen van een MPAC, hetgeen de bepaling van behandelintegriteit en contaminatie in toekomstige behandelstudies zou kunnen bevorderen.

## **Deel 5: Algemene discussie**

In het vijfde deel van dit proefschrift, de algemene discussie in hoofdstuk 9, worden allereerst de bevindingen van dit proefschrift samengevat. In het algemeen kan geconcludeerd worden dat er aanzienlijke evidentie is voor de verschillende componenten van het vrees-vermijdingsmodel, en dan met name met betrekking tot de rol van pijngerelateerde vrees in het ontwikkelen en de instandhouding van chronische klachten en het induceren van nieuwe rugpijn episoden in pijnvrije mensen. Er zijn echter enkele ambigue resultaten ten aanzien van het vrees-vermijdingsmodel. Ten eerste, de evidentie betreffende de veronderstelde disuse en fysieke deconditionering is tot nog toe meerduidelijk. Ten tweede is op dit moment de validiteit van het gehele model in het *verklaren* van CLRP klachten (nog) niet aangetoond. Tot slot wordt een belangrijke aanname van het vrees-vermijdingsmodel niet door de resultaten uit hoofdstuk 3 ondersteund, namelijk dat pijngerelateerde vrees een mediator is tussen pijn catastroferen en beperkingen. Ten aanzien van de meetmethoden voor pijngerelateerde vrees kan gesteld worden dat de TSK-G een geschikte zelfrapportage instrument is voor het bepalen van bewegingsvrees voornamelijk in pijnvrije mensen, dat de PHODA-SeV betrouwbaar en valide is voor het meten van de waargenomen schadelijkheid van activiteiten in CLRP patiënten, en dat impliciete associatie maten, althans de EAST en de GNAT, op dit moment onvoldoende betrouwbaar zijn voor toepassingen op het gebied van pijngerelateerde vrees. Ten aanzien van behandelingen wordt gesteld dat zowel EXP als GA geschikt zijn voor het verminderen van functionele beperkingen en individuele klachten. Hoewel EXP hierin minstens even effectief blijkt als GA, is deze behandeling duidelijk effectiever wanneer de doelstelling het reduceren van pijn catastroferen en vrees betreft. Dit zijn eveneens de specifieke werkingsmechanismen van EXP ten opzichte van GA. Bovendien zijn er geen aanwijzingen gevonden dat de keuze van de behandeling afgestemd moet worden op het niveau van pijngerelateerde vrees, hoewel de klinische inschatting is dat EXP niet aangewezen lijkt in absolute afwezigheid van vrees. Uit onafhankelijke oordelen van twee beoordelaars is gebleken dat zowel EXP als GA volgens protocol zijn uitgevoerd, dat wil zeggen dat de essentiële behandelingsrediënten voldoende aan bod zijn gekomen terwijl verboden behandelelementen juist afwezig waren. Vergelijkingen in effectiviteit tussen deze behandelingen lijken dus gerechtvaardigd. De bevindingen van dit proefschrift worden niet alleen geplaatst binnen de reeds bekende literatuur, maar ze worden ook kritisch beschouwd binnen de tekortkomingen van de verschillende studies, van waaruit aanbevelingen voor toekomstige studies worden gedaan. Ook worden er in dit hoofdstuk enige klinische overwegingen voor de uitvoering van EXP aangestipt en worden de implicaties van pijngerelateerde vrees voor secundaire en primaire preventie van rugklachten besproken. Tenslotte worden enkele overige aanbevelingen en aandachtspunten voor toekomstig onderzoek gegeven, waarna een slotconclusie wordt gepresenteerd.



## References





- Aldrich S, Eccleston C, Crombez G. Worrying about chronic pain: vigilance to threat and misdirected problem solving. *Behaviour Research and Therapy* 2000;38:457-470.
- Al-Obaidi SM, Al-Zoabi B, Al-Shuwaie N, Al-Zaabie N, Nelson RM. The influence of pain and pain-related fear and disability beliefs on walking velocity in chronic low back pain. *International Journal of Rehabilitation Research* 2003;26:101-108.
- Al-Obaidi SM, Beattie P, Al-Zoabi B, Al-Wekeel S. The relationship of anticipated pain and fear avoidance beliefs to outcome in patients with chronic low back pain who are not receiving workers' compensation. *Spine* 2005;30:1051-1057.
- Al-Obaidi SM, Nelson R, M., Al-Awadhi S, Al-Shuwaie N. The role of anticipation and fear of pain in the persistence of avoidance behavior in patients with chronic low back pain. *Spine* 2000;25:1126-1131.
- Arntz A, Claassens L. The meaning of pain influences its experienced intensity. *Pain* 2004;109:20-25.
- Arntz A, Dreesen L, De Jong P. The influence of anxiety on pain: Attentional and attributional mediators. *Pain* 1994;56:307-314.
- Arrindell WA, Ettema JHM. Handleiding bij een multidimensionele psychopathologie-indicator, Lisse, 1986.
- Asmundson GJ, Norton PJ, Vlaeyen JWS. Fear-avoidance models of chronic pain: an overview. In: Asmundson GJ, Vlaeyen JWS and Crombez G (Eds.). *Understanding and treating fear of pain*. New York: Oxford University Press, 2004. pp. 3-24.
- Asmundson GJ, Wright KD, Hadjistavropoulos HD. Anxiety sensitivity and disabling chronic health conditions: state of the art and future directions. *Scandinavian Journal of Behaviour Therapy* 2000;29:100-117.
- Asmundson GJG, Norton PJ, Norton GR. Beyond pain: The role of fear and avoidance in chronicity. *Clinical Psychology Review* 1999;19:97-119.
- Asmundson GJG, Wright KD, Hadjistavropoulos HD. Hypervigilance and attentional fixedness in chronic musculoskeletal pain: consistency of findings across modified stroop and dot-probe tasks. *The Journal of Pain* 2005;6:497-506.
- Banaji MR, Greenwald AG. Implicit gender stereotyping in judgments of fame. *Journal of Personality and Social Psychology* 1995;68:181-198.
- Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 1986; 51(6):1173-1182.
- Beck AT. Cognitive therapy: Nature and relation to behavior therapy. *Behaviour Therapy* 1970;1(2):184-200.
- Beck AT. *Cognitive therapy and the emotional disorders*, Meridian, New York, 1976.
- Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review* 1988;8:77-100.
- Beck AT, Ward CH, Mendelson M. An inventory for measuring depression. *Archives of General Psychiatry* 1961;4:561-571.
- Bellg AJ, Borrelli B, Resnick B, Hecht J, Minicucci DS, Ory M, Ogedegbe G, Orwig D, Ernst D, Czajkowski S. Enhancing Treatment Fidelity in Health Behavior Change Studies: Best Practices and Recommendations From the NIH Behavior Change Consortium. *Health Psychology* 2004;23:443-451.
- Beurskens AJHM, de Vet HCW, Köke AJA. Responsiveness of functional status in low back pain: A comparison of different instruments. *Pain* 1996;65:71-76.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-310.
- Bland JM, Altman DG. Statistics notes: transforming data. *BMJ* 1996;312:770.
- Bliokas VV, Cartmill TK, Nagy B. Does Systematic Graded Exposure In Vivo Enhance Outcomes in Multidisciplinary Chronic Pain Management Groups? *The Clinical Journal of Pain* 2007;23:361-374.

- Block AR, Kremer EF, Gaylor M. Behavioral treatment of chronic pain: the spouse as a discriminative cue for pain behavior. *Pain* 1980;9:243-52.
- Boersma K, Linton SJ. How does persistent pain develop? An analysis of the relationship between psychological variables, pain and function across stages of chronicity. *Behaviour Research and Therapy* 2005a;43:1495-1507.
- Boersma K, Linton SJ. Screening to identify patients at risk: profiles of psychological risk factors for early intervention. *The Clinical Journal of Pain* 2005b;21:38-43.
- Boersma K, Linton SJ, Overmeer T, Jansson M, Vlaeyen JWS, de Jong J. Lowering fear-avoidance and enhancing function through exposure in vivo; A multiple baseline study across six patients with back pain. *Pain* 2004;108:8-16.
- Borrelli B, Sepinwall D, Ernst D, Bellg AJ, Czajkowski S, Breger R, DeFrancesco C, Levesque C, Sharp DL, Ogedegbe G, Resnick B, Orwig D. A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. *Journal of Consulting and Clinical Psychology* 2005;73:852-860.
- Bosson JK, Swann WB, Pennebaker JW. Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology* 2000;79:631-643.
- Boston A, Sharpe L. The role of threat-expectancy in acute pain: effects on attentional bias, coping strategy effectiveness and response to pain. *Pain* 2005;119:168-175.
- Bousema EJ, Verbunt JA, Seelen HAM, Vlaeyen JWS, Knottnerus JA. Disuse and physical deconditioning in the first year after the onset of back pain. *Pain* 2007;130(3):279-286.
- Bouton ME. Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. *Biological Psychiatry* 2002;52:976-986.
- Brox JI, Storheim K, Holm I, Friis A, Reikeras O. Disability, pain, psychological factors and physical performance in healthy controls, patients with sub-acute and chronic low back pain: a case-control study. *Journal of Rehabilitation Medicine* 2005;37:95-99.
- Buchbinder R, Jolley D. Effects of a media campaign on back beliefs is sustained 3 years after its cessation. *Spine* 2005;30:1323-1330.
- Buchbinder R, Jolley D. Improvements in general practitioner beliefs and stated management of back pain persist 4.5 years after the cessation of a public health media campaign. *Spine* 2007;32:E156-162.
- Buchbinder R, Jolley D, Wyatt M. Volvo award winner in clinical studies: effects of a media campaign on back pain beliefs and its potential influence on management of low back pain in general practice. *Spine* 2001;26:2535-2542.
- Buer N, Linton SJ. Fear-avoidance beliefs and catastrophizing: occurrence and risk factor in back pain and ADL in the general population. *Pain* 2002;99:485-491.
- Burns JW. Anger management style and hostility: Predicting symptom-specific physiological reactivity among chronic low back pain patients. *Journal of Behavioral Medicine* 1997;20:505-522.
- Burns JW, Kubilus A, Bruehl S, Harden RN, Lofland K. Do changes in cognitive factors influence outcome following multidisciplinary treatment for chronic pain? A cross-lagged panel analysis. *Journal of Consulting and Clinical Psychology* 2003;71:81-91.
- Burton AK, McClune TD, Clarke RD, Main CJ. Long-term follow-up of patients with low back pain attending for manipulative care: outcomes and predictors. *Manual Therapy* 2004;9:30-35.
- Burton AK, Tillotson KM, Main CJ, Hollis S. Psychosocial predictors of outcome in acute and subchronic low back trouble. *Spine* 1995;20:722-728.
- Burton AK, Waddell G, Tillotson KM, Summerton N. Information and advice to patients with back pain can have a positive effect. *Spine* 1999;24:2484-2491.
- Cannon WB. Bodily changes in pain, hunger, fear and rage: an account of recent researches into the functions of emotional excitement, Appleton-Century-Crofts, New York, 1929.

- Carver CS, Scheier MF, On the self-regulation of behavior, Cambridge University Press, Cambridge, 1998.
- Catellier DJ, Hannan PJ, Murray DM, Addy CL, Conway TL, Yang S, Rice JC. Imputation of missing data when measuring physical activity by accelerometry. *Medicine & Science in Sports & Exercise* 2005;37:S555-62.
- Chambers CT, Craig KD, Bennett SM. The impact of maternal behavior on children's pain experiences: An experimental analysis. *Journal of Pediatric Psychology* 2002;27:293-301.
- Clark ME, Kori SH, Brockel J. Kinesiophobia and chronic pain: psychometric characteristics and factor analysis of the Tampa Scale. *American Pain Society Abstracts* 1996;15:77.
- Cook AJ, Brawer PA, Vowles KE. The fear-avoidance model of chronic pain: Validation and age analysis using structural equation modeling. *Pain* 2006;121:195-206.
- Cook DJ, Sackett DL, Spitzer WO. Methodologic guidelines for systematic reviews of randomized control trials in health care from the Potsdam Consultation on Meta-Analysis. *Journal of Clinical Epidemiology* 1995;48:167-71.
- Cook TD, Campbell DT, Quasi-experimentation: design & analysis issues for field settings, Rand McNally College Publishing Company, Chicago, 1979.
- Coudeyre E, Rannou F, Tubach F, Baron G, Coriat F, Brin S, Revel M, Poiraudéau S. General practitioners' fear-avoidance beliefs influence their management of patients with low back pain. *Pain* 2006;124:330-7.
- Crombez G, Eccleston C, Baeyens F, Eelen P. When somatic information threatens, catastrophic thinking enhances attentional interference. *Pain* 1998;75:187-198.
- Crombez G, Eccleston C, Baeyens F, Van Houdenhove B, Van den Broeck A. Attention to chronic pain is dependent upon pain-related fear. *Journal of Psychosomatic Research* 1999a;47:403-410.
- Crombez G, Eccleston C, Van den Broeck A, Van Houdenhove B, Goubert L. The effects of catastrophic thinking about pain on attentional interference by pain: no mediation of negative affectivity in healthy volunteers and in patients with low back pain. *Pain Research and Management* 2002a;7:31-39.
- Crombez G, Van Damme S, Eccleston C. Hypervigilance to pain: an experimental and clinical analysis. *Pain* 2005;116:4-7.
- Crombez G, Vlaeyen JW, Heuts PH, Lysens R. Pain-related fear is more disabling than pain itself: evidence on the role of pain-related fear in chronic back pain disability. *Pain* 1999b;80:329-39.
- Crombez G, Vlaeyen JWS. The Pain Catastrophizing Scale : Unpublished authorized Dutch/Flemish translation. 1996.
- Crombez G, Vlaeyen JWS, Eccleston C, Vansteenwegen D, Lysens R, Eelen P. Exposure to physical movements in low back pain patients: restricted effects of generalization. *Health Psychology* 2002b;21:573-578.
- Cunningham WA, Preacher KJ, Banaji MR. Implicit attitude measures: consistency, stability, and convergent validity. *Psychological Science* 2001;12:163-70.
- Davidson M, Keating JL. A comparison of five low back disability questionnaires: reliability and responsiveness. *Physical Therapy* 2002;82:8-24.
- Davies C, Self-discrepancy Theory and Chronic Pain, University of Leeds, Leeds, 2003.
- de Gier M, Peters ML, Vlaeyen JWS. Fear of pain, physical performance, and attentional processes in patients with fibromyalgia. *Pain* 2003;104:121-130.
- de Houwer J. The Implicit Association Test as a tool for studying dysfunctional associations in psychopathology: strength and limitations. *Journal of Behavior Therapy* 2002;33:115-133.
- de Houwer J. The Extrinsic Affective Simon Task. *Experimental Psychology* 2003a;50:77-85.

- de Houwer J. A structural analysis of indirect measures of attitudes. In: Musch J and Klauer KC (Eds.). *The psychology of evaluation: affective processes in cognition and emotion*. Mahwah: Lawrence Erlbaum Associates, Publishers, 2003b. pp. 219-244.
- de Houwer J, de Bruycker E. The identification-EAST as a valid measure of implicit attitudes toward alcohol-related stimuli. *Journal of Behavior Therapy and Experimental Psychiatry* 2007;38:133-143.
- de Houwer J, de Bruycker E. The implicit association test outperforms the extrinsic affective Simon task as an implicit measure of interindividual differences in attitudes. *British Journal of Social Psychology* 2007;46(2):401-421.
- de Jong JR, Goubert L, Vlaeyen JWS, Crombez G. Exposure in vivo bij chronische lage-rugpijn. / Exposure in vivo in chronic low back pain. *Gedragstherapie* 2002;35:49-69.
- de Jong JR, Vlaeyen JWS, Onghena P, Cuypers C, den Hollander M, Ruijgrok J. Reduction of pain-related fear in complex regional pain syndrome type I: the application of graded exposure in vivo. *Pain* 2005a;116:264-275.
- de Jong JR, Vlaeyen JWS, Onghena P, Goossens MEJB, Geilen M, Mulder H. Fear of movement/(re)injury in chronic low back pain: education of exposure in vivo as mediator to fear reduction? *The Clinical Journal of Pain* 2005b;21:9-17.
- de Jong PJ, Pasman W, Kindt M, Hout vdMA. A reaction time paradigm to assess (implicit) complaint-specific dysfunctional beliefs. *Behaviour Research and Therapy* 2001;39:101-113.
- de Jong PJ, Peters M. Implicit attitudes toward pain and injury in chronic low back pain patients. Paper presented at the conference of the European Association of Behavioural and Cognitive Therapies, September, Maastricht, The Netherlands 2002.
- Dehghani M, Sharpe L, Nicholas MK. Modification of attentional biases in chronic pain patients: a preliminary study. *European Journal of Pain* 2004;8:585-594.
- Denison E, Asenlof P, Lindberg P. Self-efficacy, fear avoidance, and pain intensity as predictors of disability in subacute and chronic musculoskeletal pain patients in primary health care. *Pain* 2004;111:245-252.
- Devilley GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behavior Therapy* 2000;31:73-86.
- DeVoe D, Gotshall R, McArthur T. Comparison of the RT3 Research Tracker and Tritrac R3D accelerometers. *Perceptual and Motor Skills* 2003;97:510-8.
- Dobson KS, Singer AR. Definitional and Practical Issues in the Assessment of Treatment Integrity. *Clinical Psychology: Science and Practice* 2005;12:384-387.
- Domjan M, *The essentials of conditioning and learning*, Thomson Wadsworth, Toronto, 2005.
- Eccleston C, Crombez G. Pain demands attention: A cognitive-affective model of the interruptive function of pain. *Psychological Bulletin* 1999;125:356-366.
- Eccleston C, Crombez G, Aldrich S, Stannard C. Attention and somatic awareness in chronic pain. *Pain* 1997;72:209-215.
- Edwards JR, Lambert LS. Methods for integrating moderation and mediation: a general analytical framework using moderated path analysis. *Psychological Methods* 2007;12:1-22.
- Edwards RR, Fillingim RB, Maixner W, Sigurdsson A, Haythornthwaite J. Catastrophizing predicts changes in thermal pain responses after resolution of acute dental pain. *The Journal of Pain* 2004;5:164-170.
- Eisenberger NI, Jarcho JM, Lieberman MD, Naliboff BD. An experimental study of shared sensitivity to physical pain and social rejection. *Pain* 2006;126:132-138.
- Ellis A, *Reason and emotion in psychotherapy*, Lyle-Stuart, New York, 1962.
- Ellwart T, Becker ES, Rinck M. Activation and measurement of threat associations in fear of spiders: An application of the extrinsic affective Simon task. *Journal of Behavior Therapy and Experimental Psychiatry* 2005;36:281-299.

- Essink Bot ML, Krabbe PF, Bonsel GJ, Aaronson NK. An empirical comparison of four generic health status measures. The Nottingham Health Profile, the Medical Outcomes Study 36-item Short-Form Health Survey, the COOP/WONCA charts, and the EuroQol instrument. *Medical Care* 1997;35:522-537.
- Eysenck MW, *Anxiety: the cognitive perspective*, Erlbaum, Hove, 1992.
- Faas A, Chavannes AW, Koes AW, J.M.M. VdH, J.M.A. M, I. S, A.C.M. R, Van der Laan JR. NHG-standaard lage-rugpijn [Practice Guideline 'Low Back Pain']. *Huisarts & Wetenschap* 1996;39:18-31.
- Fazio RH, Jackson JR, Dunton BC, Williams CJ. Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology* 1995;69:1013-1027.
- Fazio RH, Olson MA. Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology* 2003;54:297-327.
- Flor H, Birbaumer N. Acquisition of chronic pain. Psychophysiologic mechanisms. *APS Journal* 1994;3:119-127.
- Flor H, Birbaumer N, Schugens MM, Lutzenberger W. Symptom-specific psychophysiological responses in chronic pain patients. *Psychophysiology* 1992;29:452-460.
- Flor H, Birbaumer N, Schulte W, Roos R. Stress-related electromyographic responses in patients with chronic temporomandibular pain. *Pain* 1991;46:145-152.
- Flor H, Breitenstein C, Birbaumer N, Furst M. A psychophysiological analysis of spouse solicitousness towards pain behaviors, spouse interaction, and pain perception. *Behavior Therapy* 1995;26:255-272.
- Flor H, Kerns RD, Turk DC. The role of spouse reinforcement, perceived pain, and activity levels of chronic pain patients. *Journal of Psychosomatic Research* 1987;31:251-9.
- Flor H, Knost B, Birbaumer N. The role of operant conditioning in chronic pain: an experimental investigation. *Pain* 2002;95:111-8.
- Flor H, Turk DC, Birbaumer N. Assessment of stress-related psychophysiological reactions in chronic back pain patients. *Journal of Consulting and Clinical Psychology* 1985;53:354-364.
- Fordyce WE, *Behavioral methods for chronic pain and illness.*, Mosby, St. Louis., 1976.
- Fordyce WE. Learned pain: pain as behaviour. In: Loeser JD, Butler SH, Chapman CR and Turk DC (Eds.). *Bonica's Management of Pain*. Lippincott: Williams & Wilkins, 2001a.
- Fordyce WE. Operant or contingency therapies. In: Loeser JD, Butler SH, Chapman CR and Turk DC (Eds.). *Bonica's Management of Pain*. Lippincott: Williams & Wilkins, 2001b.
- Fordyce WE, Fowler RS, Delateur B. An Application of Behavior Modification Technique to a Problem of Chronic Pain. *Behaviour Research and Therapy* 1968.
- Fritz JM, George SZ, Delitto A. The role of fear-avoidance beliefs in acute low back pain: relationships with current and future disability and work status. *Pain* 2001;94:7-15.
- Fritz JM, Irrgang JJ. A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. *Physical Therapy* 2001;81:776-788.
- Geilen MJ, Leeuw M, Hodiament-Joosten MCJ. Graded activity volgens gedragsgeoriënteerde principes. In: Dijkstra PU, van Burken P, Marinus J, Nijs J and van Wilgen CP (Eds.). *Jaarboek Fysiotherapie Kinesitherapie*. Houten: Bohn Stafleu Van Loghum, 2005.
- Geisser ME, Haig AJ, Theisen ME. Activity avoidance and function in persons with chronic back pain. *Journal of Occupational Rehabilitation* 2000;10:215-227.
- Geisser ME, Haig AJ, Wallbom AS, Wiggert EA. Pain-related fear, lumbar flexion, and dynamic EMG among persons with chronic musculoskeletal low back pain. *The Clinical Journal of Pain* 2004;20:61-69.
- George SZ, Bialosky JE, Donald DA. The centralization phenomenon and fear-avoidance beliefs as prognostic factors for acute low back pain: a preliminary investigation involving

- patients classified for specific exercise. *Journal of Orthopaedic & Sports Physical Therapy* 2005;35:580-588.
- George SZ, Fritz JM, Bialosky JE, Donald DA. The effect of a fear-avoidance-based physical therapy intervention for patients with acute low back pain: results of a randomized clinical trial. *Spine* 2003;28:2551-2560.
- Geraets JJXR, de Groot IJM, Goossens MEJB, de Bruijn CPC, de Bie RA, van den Heuvel WJA, Dinant GJ. Comparison of two recruitment strategies for patients with chronic shoulder complaints. *British Journal of General Practice* 2006;56:127-133.
- Gheldof E, de Jong PJ, Vinck J, Houben RMA. Attitudes towards physical activity: the role of implicit versus explicit associations. In: Asmundson GJ, Vlaeyen JWS and Crombez G (Eds.). *Understanding and treating fear of pain*. New York: Oxford University Press, 2004. pp. 117-137.
- Gheldof EL, Vinck J, Vlaeyen JWS, Hidding A, Crombez G. The differential role of pain, work characteristics and pain-related fear in explaining back pain and sick leave in occupational settings. *Pain* 2005;113:71-81.
- Gheldof ELM, Vinck J, Van den Bussche E, Vlaeyen JWS, Hidding A, Crombez G. Pain and pain-related fear are associated with functional and social disability in an occupational setting: Evidence of mediation by pain-related fear. *European Journal of Pain* 2006;10:513-25.
- Gommans IHB, Koes BW, van Tulder MW. Validiteit en responsiviteit van de Nederlandstalige Roland Disability Questionnaire. *Vragenlijst naar functionele status bij patiënten met lage rugpijn*. *Nederlands Tijdschrift voor Fysiotherapie* 1997;107:28-33.
- Goossens MEJB. Economic aspects of chronic musculoskeletal pain. In: Linton SJ (Eds.). *Pain Research and Clinical Management Series*, volume 12: new avenues for the prevention of chronic musculoskeletal pain. Amsterdam: Elsevier, 2002.
- Goossens MEJB, Vlaeyen JWS, Hidding A, Kole-Snijders A, Evers SMAA. Treatment Expectancy Affects the Outcome of Cognitive-Behavioral Interventions in Chronic Pain. *The Clinical Journal of Pain* 2005;21:18-26.
- Goubert L, Crombez G, Danneels L. The reluctance to generalize corrective experiences in chronic low back pain patients: a questionnaire study of dysfunctional cognitions. *Behaviour Research and Therapy* 2005a;43:1055-1067.
- Goubert L, Crombez G, De bourdeaudhuij I. Low back pain, disability and back pain myths in a community sample: prevalence and interrelationships. *European Journal of Pain* 2004a;8:385-394.
- Goubert L, Crombez G, Hermans D, Vanderstraeten G. Implicit attitudes towards pictures of back-stressing activities in pain-free subjects and patients with low back pain: an affective priming study. *European Journal of Pain* 2003;7:33-42.
- Goubert L, Crombez G, Lysens R. Effects of varied-stimulus exposure on overpredictions of pain and behavioural performance in low back pain patients. *Behaviour Research and Therapy* 2005b;43:1347-1361.
- Goubert L, Crombez G, Peters M. Pain-related fear and avoidance: a conditioning perspective. In: Asmundson GJG, Vlaeyen JWS and Crombez G (Eds.). *Understanding and treating fear of pain*. New York: Oxford University Press, 2004b. pp. 25-50.
- Goubert L, Crombez G, Van Damme S. The role of neuroticism, pain catastrophizing and pain-related fear in vigilance to pain: a structural equations approach. *Pain* 2004c;107:234-241.
- Goubert L, Crombez G, Van Damme S, Vlaeyen JWS, Bijttebier P, Roelofs J. Confirmatory factor analysis of the Tampa Scale for Kinesiophobia: invariant two-factor model across low back pain patients and fibromyalgia patients. *The Clinical Journal of Pain* 2004d;20:103-110.

- Goubert L, Crombez G, Vlaeyen JWS, Van Damme S, van den Broeck A, van Houdenhove B. De Tampa schaal voor kinesiofobie: Psychometrische karakteristieken en normering. *Gedrag en Gezondheid* 2000;28:54-62.
- Goubert L, Francken G, Crombez G, Vansteenwegen D, Lysens R. Exposure to physical movement in chronic back pain patients: No evidence for generalization across different movements. *Behaviour Research and Therapy* 2002;40:415-429.
- Granot M, Ferber SG. The roles of pain catastrophizing and anxiety in the prediction of postoperative pain intensity: a prospective study. *The Clinical Journal of Pain* 2005;21:439-445.
- Green LW, Glasgow RE. Evaluating the relevance, generalization, and applicability of research: issues in external validation and translation methodology. *Evaluation & The Health Professions* 2006;29:126-53.
- Greenwald AG, McGhee DE, Schwartz JLK. Measuring individual differences in implicit cognition: the Implicit Association Test. *Journal of Personality and Social Psychology* 1998;14:1464-1480.
- Greenwald AG, Nosek BA. Health of the Implicit Association Test at age 3. *Zeitschrift fur Experimentelle Psychologie* 2001;48:85-93.
- Gresham FM. Treatment Integrity and Therapeutic Change: Commentary on Perepletchikova and Kazdin. *Clinical Psychology: Science and Practice* 2005;12:391-394.
- Gresham FM, Gansle KA, Noell GH. Treatment integrity in applied behavior analysis with children. *Journal of Applied Behavior Analysis* 1993;26:257-263.
- Gretebeck RJ, Montoye HJ. Variability of some objective measures of physical activity. *Medicine & Science in Sports & Exercise* 1992;24:1167-72.
- Grotle M, Brox JI, Vollestad NK. Concurrent comparison of responsiveness in pain and functional status measurements used for patients with low back pain. *Spine* 2004a;29:E492-501.
- Grotle M, Vollestad NK, Veierod MB, Brox JI. Fear-avoidance beliefs and distress in relation to disability in acute and chronic low back pain. *Pain* 2004b;112:343-352.
- Guzman J, Esmail R, Karjalainen K, Malmivaara A, Irvin E, Bombardier C. Multidisciplinary rehabilitation for chronic low back pain: systematic review. *BMJ* 2001;322:1511 - 1516.
- Hadjistavropoulos HD, Kowalyk KM. Patient-therapist relationships among patients with pain-related fear. In: Asmundson GJG, Vlaeyen JWS and Combez G (Eds.). *Understanding and treating fear of pain*. New York: Oxford University Press, 2004. pp. 237-264.
- Hanley JA, McNiell BJ. The meaning and use of the area under the Receiver Operator Characteristic (ROC) curve. *Radiology* 1982;143:29-36.
- Hanley MA, Jensen MP, Ehde DM, Hoffman AJ, Patterson DR, Robinson LR. Psychosocial predictors of long-term adjustment to lower-limb amputation and phantom limb pain. *Disability and Rehabilitation* 2004;26:882-893.
- Haythornthwaite JA, Clark MR, Pappagallo M, Raja SN. Pain coping strategies play a role in the persistence of pain in post-herpetic neuralgia. *Pain* 2003;106:453-460.
- Hermans D, Craske MG, Mineka S, Lovibond PF. Extinction in human fear conditioning. *Biological Psychiatry* 2006;30:361-368.
- Heuts PHTG, Vlaeyen JWS, Roelofs J, de Bie RA, Aretz K, van Weel C, van Schayck OC. Pain-related fear and daily functioning in patients with osteoarthritis. *Pain* 2004;110:228-235.
- Hicks GE, Fritz JM, Delitto A, McGill SM. Preliminary development of a clinical prediction rule for determining which patients with low back pain will respond to a stabilization exercise program. *Archives Of Physical Medicine and Rehabilitation* 2005;86:1753-1762.
- Hoffman BM, Papas RK, Chatkoff DK, Kerns RD. Meta-analysis of psychological interventions for chronic low back pain. *Health Psychology* 2007;26:1-9.



- Hofmann W, Gawronski B, SGschwendner T, Le H, Schmitt M. A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin* 2005;31:1369-1385.
- Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ* 1999;319:670-674.
- Holzl R, Kleinbohl D, Huse E. Implicit operant learning of pain sensitization. *Pain* 2005;115:12-20.
- Horvath AO, Symonds BD. Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology* 1991;38:139-149.
- Houben RMA, Gijzen A, Peterson J, de Jong PJ, Vlaeyen JWS. Do health care providers' attitudes towards back pain predict their treatment recommendations? Differential predictive validity of implicit and explicit attitude measures. *Pain* 2005a;114:491-498.
- Houben RMA, Leeuw M, Vlaeyen JWS, Goubert L, Picavet HSJ. Fear of movement/injury in the general population: Factor structure and psychometric properties of an adapted version of the Tampa Scale for Kinesiophobia. *Journal of Behavioural Medicine* 2005b;28:415-424.
- Houben RMA, Ostelo RWJG, Vlaeyen JWS, Wolters PMJC, Peters M, Stomp-van den Berg SGM. Health care providers' orientations towards common low back pain predict perceived harmfulness of physical activities and recommendations regarding return to normal activity. *European Journal of Pain* 2005c;9:173-183.
- Houben RMA, Vlaeyen JWS, Peters M, Ostelo RWJG, Wolters PMJC, Stomp-van den Berg SGM. Health care providers' attitudes and beliefs towards common low back pain: factor structure and psychometric properties of the HC-PAIRS. *The Clinical Journal of Pain* 2004;20:37-44.
- Huijding J, de Jong PJ. A Pictorial Version of the Extrinsic Affective Simon Task: Sensitivity to Generally Affective and Phobia-Relevant Stimuli in High and Low Spider Fearful Individuals. *Experimental Psychology* 2005;52:289-295.
- Huijding J, de Jong PJ. Specific predictive power of automatic spider-related affective associations for controllable and uncontrollable fear responses toward spiders. *Behaviour Research and Therapy* 2006;44:161-176.
- Jellema P, van der Horst HE, Vlaeyen JWS, Stalman WAB, Bouter LM, van der Windt DAWM. Predictors of outcome in patients with (sub)acute low back pain differ across treatment groups. *Spine* 2006;31:1699-1705.
- Jensen MP, Turner JA, Romano JM. Changes in beliefs, catastrophizing, and coping are associated with improvement in multidisciplinary pain treatment. *Journal of Consulting and Clinical Psychology* 2001;69:655-662.
- Jolliffe CD, Nicholas MK. Verbally reinforcing pain reports: an experimental test of the operant model of chronic pain. *Pain* 2004;107:167-75.
- Joreskog KG, Sorbom D, Windows LISREL version 8.54, Scientific Software International, Chicago, 2001.
- Kanfer FH, Philips JS, Learning foundations of behaviour therapy, Wiley, New York, 1970.
- Karoly P, Ruehlman LS. Motivational implications of pain: chronicity, psychological distress, and work goal construal in a national sample of adults. *Health Psychology* 1996;15:383-390.
- Kazdin AE. Comparative outcome studies of psychotherapy: Methodological issues and strategies. *Journal of Consulting and Clinical Psychology* 1986;54:95-105.
- Kazdin AE, Nock MK. Delineating mechanisms of change in child and adolescent therapy: methodological issues and research recommendations. *The Journal of Child Psychology and Psychiatry* 2003;44:1116-29.
- Keogh E, Asmundson GJG. Negative affectivity, catastrophizing, and anxiety sensitivity. In: Asmundson GJG, Vlaeyen JWS and Crombez G (Eds.). *Understanding and treating fear of pain*. New York: Oxford University Press, 2004. pp. 91-116.

- Keogh E, Ellery D, Hunt C, Hannent I. Selective attentional bias for pain-related stimuli amongst pain fearful individuals. *Pain* 2001;91:91-100.
- Kerns RD, Bayer LA, Findley JC. Motivation and adherence in the management of chronic pain. In: Block AR, Fernandez E and Kremer EF (Eds.). *Handbook of Pain Syndromes: Biopsychosocial Perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates, 1999. pp. 99-121.
- Klaber Moffett JA, Carr J, Howarth EH. High fear-avoiders of physical activity benefit from an exercise program for patients with back pain. *Spine* 2004;29:1167-1179.
- Klenerman L, Slade PD, Stanley IM, Pennie B, Reilly JP, Atchison LE, Troup JD, Rose MJ. The prediction of chronicity in patients with an acute attack of low back pain in a general practice setting. *Spine* 1995;20:478-84.
- Koes BW, van Tulder MW, Ostelo R, Kim Burton A, Waddell G. Clinical guidelines for the management of low back pain in primary care: an international comparison. *Spine* 2001;26:2504-2513.
- Kole-Snijders AM, Vlaeyen JW, Goossens ME, Rutten-van Molken MP, Heuts PH, van Breukelen G, van Eek H. Chronic low-back pain: what does cognitive coping skills training add to operant behavioral treatment? Results of a randomized clinical trial. *Journal of Consulting and Clinical Psychology* 1999;67:931-44.
- Kopec JA, Esdaile JM, Abrahamowicz M, Abenheim L, Wood Dauphinee S, Lamping DL, Williams JI. The Quebec Back Pain Disability Scale: conceptualization and development. *Journal of Clinical Epidemiology* 1996;49:151-61.
- Kopec JA, Esdaile JM, Abrahamowicz M, Abenheim L, Wood-Dauphinee S, Lamping DL, Williams JI. The Quebec Back Pain Disability Scale: measurement properties. *Spine* 1995;20:341-352.
- Kori SH, Miller RP, Todd DD. Kinesiophobia: A new view of chronic pain behavior. *Pain Management* 1990;Jan/Feb:35-43.
- Kramer M, Ebert V, Kinzl L, Dehner C, Elbel M, Hartwig E. Surface electromyography of the paravertebral muscles in patients with chronic low back pain. *Archives of Physical Medicine and Rehabilitation* 2005;86:31-36.
- Kugler K, Wijn J, Geilen M, de Jong J, Vlaeyen JWS, The Photograph series of Daily Activities (PHODA). CD-rom version 1.0., Institute for Rehabilitation Research and School for Physiotherapy Heerlen, The Netherlands, 1999.
- Kvist J, Ek A, Sporrstedt K, Good L. Fear of re-injury: a hindrance for returning to sports after anterior cruciate ligament reconstruction. *Knee Surgery Sports Traumatology Arthroscopy* 2005;13:393-397.
- Laird NM, Wang F. Estimating rates of changes in randomized clinical trials. *Controlled Clinical Trials* 1990;11:405-419.
- Lamoth CJ, Daffertshofer A, Meijer OG, Lorimer Moseley G, Wuisman PI, Beek PJ. Effects of experimentally induced pain and fear of pain on trunk coordination and back muscle activity during walking. *Clinical Biomechanics (Bristol, Avon)* 2004;19:551-563.
- Lang PJ. A bio-informational theory of emotional imagery. *Psychophysiology* 1979;16:495-512.
- Leeuw M, Goossens MEJB, Linton SJ, Crombez G, Boersma K, Vlaeyen JWS. The fear avoidance model of musculoskeletal pain: current state of scientific evidence. *Journal of Behavioral Medicine* 2007a;30:77-94.
- Leeuw M, Goossens MEJB, van Breukelen GJP, Boersma K, Vlaeyen JWS. Measuring perceived harmfulness of physical activities in patients with chronic low back pain: the Photograph Series of Daily Activities - Short electronic Version. *The Journal of Pain* 2007b; 8(11):840-849.
- Leeuw M, Houben RMA, Severeijns R, Picavet HSJ, Schouten EGW, Vlaeyen JWS. Pain-related fear in low back pain: a prospective study in the general population. *European Journal of Pain* 2007c;11:256-266.

- Leeuw M, Vlaeyen JWS, de Jong JR, Goossens MEJB, Exposure in vivo bij chronische lage rugpijn, werkboek en therapeutenboek, Boom Test Uitgevers, 2006.
- Lefair Mj, Watt MC, Stewart SH, Wright KD. Implicit associations between anxiety-related symptoms and catastrophic consequences in high anxiety sensitive individuals. *Cognition and Emotion* 2006;20:295-308.
- Lethem J, Slade PD, Troup JD, Bentley G. Outline of a fear-avoidance model of exaggerated pain perception: I. *Behaviour Research and Therapy* 1983;21:401-408.
- Lichstein KL, Riedel BW, Grieve R. Fair tests of clinical trials: A treatment implementation model. *Advances in Behaviour Research and Therapy* 1994;16:1-29.
- Lilienfeld SO, Jacob RG, Turner SM. Anxiety sensitivity: An examination of theoretical and methodological issues. *Advances in Behaviour Research and Therapy* 1993;15:147-183.
- Lindstrom I, Ohlund C, Eek C, Wallin L, Peterson LE, Fordyce WE, Nachemson AL. The effect of graded activity on patients with subacute low back pain: a randomized prospective clinical study with an operant-conditioning behavioral approach. *Physical Therapy* 1992;72:279-90; discussion 291-3.
- Linton SJ. A review of psychological risk factors in back and neck pain. *Spine* 2000;25:1148-1156.
- Linton SJ. Do psychological factors increase the risk for back pain in the general population in both a cross-sectional and prospective analysis? *Pain* 2005;9:355-361.
- Linton SJ, Boersma K. Early identification of patients at risk of developing a persistent back problem: the predictive validity of the Orebro Musculoskeletal Pain Questionnaire. *The Clinical Journal of Pain* 2003;19:80-86.
- Linton SJ, Boersma K, Jansson M, Svard L, Botvalde M. The effects of cognitive-behavioral and physical therapy preventive interventions on pain-related sick leave: a randomized controlled trial. *The Clinical Journal of Pain* 2005;21:109-119.
- Linton SJ, Buer N, Vlaeyen JWS, Hellsing A-L. Are fear-avoidance beliefs related to the inception of an episode of back pain? A prospective study. *Psychology and Health* 2000;14:1051-1059.
- Linton SJ, Gotestam KG. Controlling pain reports through operant conditioning: A laboratory demonstration. *Perceptual and Motor Skills* 1985;60:427-437.
- Linton SJ, Hallden K. Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and subacute back pain. *The Clinical Journal of Pain* 1998;14:209-215.
- Linton SJ, Nordin E. A 5-year follow-up evaluation of the health and economic consequences of an early cognitive behavioral intervention for back pain: a randomized, controlled trial. *Spine* 2006;31:853-858.
- Linton SJ, Overmeer T, Janson M, Vlaeyen JWS, de Jong JR. Graded in-vivo exposure treatment for fear-avoidant pain patients with functional disability: A case study. *Cognitive Behaviour Therapy* 2002a;31:49-58.
- Linton SJ, Ryberg M. Do epidemiological results replicate? The prevalence and health-economic consequences of neck and back pain in the general population. *European Journal of Pain* 2000;4:347-54.
- Linton SJ, Ryberg M. A cognitive-behavioral group intervention as prevention for persistent neck and back pain in a non-patient population: a randomized controlled trial. *Pain* 2001;90.
- Linton SJ, Vlaeyen JWS, Ostelo RWJG. The back pain beliefs of health care providers: are we fear-avoidant? *Journal of Occupational Rehabilitation* 2002b;12:223-232.
- Little RJA. Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association* 1995;90:1112-1121.
- Loughnan S, Haslam N. Animals and Androids: Implicit Associations Between Social Categories and Nonhumans. *Psychological Science* 2007;18:116-121.

- Lousberg R, Schmidt AJ, Groenman NH. The relationship between spouse solicitude and pain behavior: searching for more experimental evidence. *Pain* 1992;51:75-9.
- Luborsky L, DeRubeis RJ. The use of psychotherapy treatment manuals: A small revolution in psychotherapy research style. *Clinical Psychology Review* 1984;4:5-14.
- MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods* 2002;7:83-104.
- Maetzel A, Li L. The economic burden of low back pain: a review of studies published between 1996 and 2001. *Best Practice & Research Clinical Rheumatology* 2002;16:23-30.
- Mallinckrodt B, Abraham WT, Wei M, Russell DW. Advances in Testing the Statistical Significance of Mediation Effects. *Journal of Counseling Psychology* 2006;53:372-378.
- Mannion AF, Junge A, Taimela S, Muntener M, Lorenzo K, Dvorak J. Active therapy for chronic low back pain: part 3. Factors influencing self-rated disability and its change following therapy. *Spine* 2001;26:920-929.
- Mannion AF, Muntener M, Taimela S, Dvorak J. A randomized clinical trial of three active therapies for chronic low back pain. *Spine* 1999;24:2435-48.
- Masse LC, Fuemmeler BF, Anderson CB, Matthews CE, Trost SG, Catellier DJ, Treuth M. Accelerometer data reduction: a comparison of four reduction algorithms on select outcome variables. *Medicine & Science in Sports & Exercise* 2005;37:S544-54.
- Mayo-Wilson E. Reporting Implementation in Randomized Trials: Proposed Additions to the Consolidated Standards of Reporting Trials Statement. *American Journal of Public Health* 2007;97:630-633.
- McCracken LM, Gross RT, Eccleston C. Multimethod assessment of treatment process in chronic low back pain: comparison of reported pain-related anxiety with directly measured physical capacity. *Behaviour Research and Therapy* 2002;40:585-594.
- McCracken LM, Zayfert C, Gross RT. The Pain Anxiety Symptoms Scale: development and validation of a scale to measure fear of pain. *Pain* 1992;50:67-73.
- McNeil DW, Rainwater AJ. Development of the fear of pain questionnaire-III. *Journal of Behavioral Medicine* 1998;21:389-411.
- McNeil DW, Vowles KE. Assessment of fear and anxiety associated with pain: conceptualization, methods, and measures. In: Asmundson GJG, Vlaeyen JWS and Crombez G (Eds.). *Understanding and treating fear of pain*. New York: Oxford University Press, 2004. pp. 189-211.
- Mead N, Bower P. Patient-centredness: A conceptual framework and review of the empirical literature. *Social Science and Medicine* 2000;51:1087-1110.
- Meerding WJ, Bonneux L, Polder JJ, Koopmanschap MA, van der Maas PJ. Demographic and epidemiological determinants of healthcare costs in Netherlands: cost of illness study. *British Medical Journal* 1998;317.
- Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975;1:277-99.
- Miller RP, Kori SH, Todd DD. The Tampa Scale for Kinesophobia: Unpublished report, Tampa, FL, 1991.
- Mineka S, Zinbarg R. A contemporary learning theory perspective on the etiology of anxiety disorders: it's not what you thought it was. *American Psychologist* 2006;61:10-26.
- Moerbeek M, van Breukelen GJ, Berger MP. A comparison between traditional methods and multilevel regression for the analysis of multicenter intervention studies. *Journal of Clinical Epidemiology* 2003;56:341-50.
- Moher D, Schulz KF, Altman D. The CONSORT Statement: revised recommendations for improving the quality of reports of parallel-group randomized trials 2001. *Explore (NY)* 2005;1:40-5.
- Moncher FJ, Prinz RJ. Treatment fidelity in outcome studies. *Clinical Psychology Review* 1991;11:247-266.

- Moore JE, BVon Korff M, Cherkin D, Saunders K, Lorig K. A randomized trial of a cognitive-behavioral program for enhancing back pain self care in a primary care setting. *Pain* 2000;88:145-153.
- Morley S. Single case methodology in psychological therapy. In: Lindsey SJE and Powell GE (Eds.). *The handbook of clinical adult psychology* (3rd ed.). London: Psychology Press and Brunner-Routledge, In press.
- Morley S, Eccleston C. The object of fear in pain. In: Asmundson GJ, Vlaeyen JWS and Crombez G (Eds.). *Understanding and treating fear of pain*. New York: Oxford University Press, 2004. pp. 163-188.
- Morley S, Eccleston C, Williams A. Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. *Pain* 1999;80:1-13.
- Morley S, Williams ACdC. Conducting and evaluating treatment outcome studies. In: Turk DC and Gatchel RJ (Eds.). *Psychological Approaches to Pain Management*. New York: The Guilford press, 2002. pp. 52-68.
- Morley S, Williams ACdC, Black S. A confirmatory factor analysis of the Beck Depression Inventory in chronic pain. *Pain* 2002;99:289-298.
- Mousavi SJ, Parnianpour M, Mehdian H, Montazeri A, Mobini B. The Oswestry Disability Index, the Roland-Morris Disability Questionnaire, and the Quebec Back Pain Disability Scale: translation and validation studies of the Iranian versions. *Spine* 2006;31:E454-E459.
- Mowrer OH. On the dual nature of learning: a reinterpretation of 'conditioning' and 'problem-solving'. *Harvard Educational Review* 1947;17:102-150.
- Nachemson AL. Newest knowledge of low back pain. A critical look. *Clinical Orthopaedics And Related Research* 1992;279:8-20.
- Nachemson AL, Vingard E. Assessment of patients with neck and back pain: a best-evidence synthesis. In: Nachemson AL and Johnsson E (Eds.). *Neck and back pain: the scientific evidence of causes, diagnosis, and treatment*. Philadelphia: Lippincott Williams & Wilkins, 2000. pp. 189-235.
- Nachemson AL, Waddell G, Norlund AL. Epidemiology of neck and low back pain. In: Nachemson AL and Johnsson E (Eds.). *Neck and back pain; the scientific evidence of causes, diagnosis, and treatment*. Philadelphia: Lippincott Williams & Wilkins, 2000. pp. 165-187.
- Nash JM, D.M. W, Nicholson R, Trask PC. The contribution of pain-related anxiety to disability from headache. *Journal of Behavioural Medicine* 2006;29:61-67.
- Nederhand MJ, Ijzerman MJ, Hermens HJ, Turk DC, Zilvold G. Predictive value of fear avoidance in developing chronic neck pain disability: consequences for clinical decision making. *Archives of Physical Medicine and Rehabilitation* 2004;85:496-501.
- Nielens H, Plaghki L. Cardiorespiratory fitness, physical activity level, and chronic pain: are men more affected than women? *The Clinical Journal of Pain* 2001;17:129-137.
- Nijs J, Vanherberghen K, Duquet W, De Meirleir K. Chronic fatigue syndrome: lack of association between pain-related fear of movement and exercise capacity and disability. *Physical Therapy* 2004;84:696-705.
- Norlund AI, Waddell G. Cost of back pain in some OECD countries. In: Nachemson AL and Johnson E (Eds.). *Neck and back pain: the scientific evidence of causes, diagnosis, and treatment*. Philadelphia: Lippincott Williams & Wilkins, 2000. pp. 421-426.
- Norton PJ, Asmundson GJG. Anxiety sensitivity, fear, and avoidance behavior in headache pain. *Pain* 2004;111:218-223.
- Nosek BA, Banaji MR. The Go/No-Go association Task. *Social Cognition* 2001;19:625-664.
- Onghena P, Edgington ES. Customization of pain treatments: single-case design and analysis. *The Clinical Journal of Pain* 2005;21:56-68.

- Osman A, Barrios FX, Gutierrez PM, Kopper BA, Merrifield T, Grittmann L. The Pain Catastrophizing Scale: Further psychometric evaluation with adult samples. *Journal of Behavioral Medicine* 2000;23:351-365.
- Osman A, Barrios FX, Kopper BA, Hauptmann W, Jones J, O'Neill E. Factor structure, reliability and validity of the pain catastrophizing scale. *Journal of Behavioral Medicine* 1997;20:589-605.
- Ostelo RW, van Tulder MW, Vlaeyen JWS, Linton SJ, Morley SJ, Assendelft WJ. Behavioural treatment for chronic low-back pain. *Cochrane Database Syst Rev* 2005;Cd002014.
- Pavlin DJ, Sullivan MJ, Freund PR, Roesen K. Catastrophizing: a risk factor for postsurgical pain. *The Clinical Journal of Pain* 2005;21:83-90.
- Perepletchikova F, Kazdin AE. Treatment Integrity and Therapeutic Change: Issues and Research Recommendations. *Clinical Psychology: Science and Practice* 2005;12:365-383.
- Peters ML, Schmidt AJ. Psychophysiological responses to repeated acute pain stimulation in chronic low back pain patients. *Journal of Psychosomatic Research* 1991;35:59-74.
- Peters ML, Vlaeyen JWS, Kunnen AMW. Is pain-related fear a predictor of somatosensory hypervigilance in chronic low back pain patients? *Behaviour Research and Therapy* 2002;40:85-103.
- Peters ML, Vlaeyen JWS, van Drunen C. Do fibromyalgia patients display hypervigilance for innocuous somatosensory stimuli? Application of a body scanning reaction time paradigm. *Pain* 2000;86:283-292.
- Peters ML, Vlaeyen JWS, Weber WEJ. The joint contribution of physical pathology, pain-related fear and catastrophizing to chronic back pain disability. *Pain* 2005;113:45-50.
- Peterson L, Homer AL, Wonderlich SA. The integrity of independent variables in behavior analysis. *Journal of Applied Behavior Analysis* 1982;15:477-92.
- Pfingsten M, Leibing E, Harter W, Kröner-Herwig B, Hempel D, Kronshage U, Hildebrandt J. Fear-avoidance behavior and anticipation of pain in patients with chronic low back pain: A randomized controlled study. *Pain Medicine* 2001;2:259-266.
- Philips HC. Avoidance behaviour and its role in sustaining chronic pain. *Behaviour Research and Therapy* 1987;25:273-279.
- Picavet HS, Hazes JM. Prevalence of self reported musculoskeletal diseases is high. *Annals of the Rheumatic Diseases* 2003.
- Picavet HS, Hoeymans N. Health related quality of life in multiple musculoskeletal diseases: SF-36 and EQ-5D in the DMC3 study. *Annals of the Rheumatic Diseases* 2004;63(6):723-729.
- Picavet HS, Schouten JS, Smit HA. Prevalence and consequences of low back problems in The Netherlands, working vs non-working population, the MORGEN-Study. Monitoring Project on Risk Factors for Chronic Disease. *Public Health* 1999;113:73-7.
- Picavet HS, van den Bos GA. The contribution of six chronic conditions to the total burden of mobility disability in the Dutch population. *American Journal of Public Health* 1997;87:1680-2.
- Picavet HS, Vlaeyen JWS, Schouten JS. Pain catastrophizing and kinesiophobia: predictors of chronic low back pain. *American Journal of Epidemiology* 2002;156:1028-1034.
- Picavet HSI, Schouten JSAG. Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC(3)-study. *Pain* 2003;102:167-178.
- Picavet HSI, van Gils HWV, Schouten JSAG, Klachten van het bewegingsapparaat in de Nederlandse bevolking: prevalenties, consequenties en risicogroepen (in Dutch) (Musculoskeletal complaints in the Dutch population: prevalence, consequences and risk groups), CBS/RIVM (RIVM rapportnummer: 266807002), Bilthoven, the Netherlands, 2000.

- Pincus T, Burton AK, Vogel S, Field AP. A systematic review of psychological factors as predictors of chronicity/disability in prospective cohorts of low back. *Spine* 2002;27:E109-120.
- Pincus T, Morley S. Cognitive-processing bias in chronic pain: a review and integration. *Psychological Bulletin* 2001;127:599-617.
- Pituch KA, Whittaker TA, Stapleton LM. A Comparison of Methods to Test for Mediation in Multisite Experiments. *Multivariate Behavioral Research* 2005;40:1-23.
- Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, Gaboury I. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *The Medical Journal of Australia* 2006;185:263-7.
- Poiraudau S, Rannou F, Le Henanff A, Coudeyre E, Rozenberg S, Huas D, Martineau C, Jolivet-Landreau I, Revel M, Ravaud P. Outcome of subacute low back pain: influence of patients' and rheumatologists' characteristics. *Rheumatology-(Oxford)* 2006;45:718-723.
- Powell SM, Jones DI, Rowlands AV. Technical variability of the RT3 accelerometer. *Medicine & Science in Sports & Exercise* 2003;35:1773-8.
- Preacher KJ, Leonardelli GJ. Calculation for the Sobel test: an interactive calculation tool for mediation [computer software]. Available from: <http://www.psych.ku.edu/preacher/sobel/sobel.htm>, 2001.
- Psychology Software Tools I, Serial Response Box, 1996-2003.
- Rachman S. Anxiety, Psychological Press, Hove, 1998.
- Rachman S. Fear and Courage: A Psychological Perspective. *Social Research* 2004;71:149-176.
- Rachman S, Hodgson R. Synchrony and desynchrony in fear and avoidance. *Behaviour Research and Therapy* 1974;12:311-318.
- Rainville J, Carlson N, Polatin P, Gatchel RJ, Indahl A. Exploration of physicians' recommendations for activities in chronic low back pain. *Spine* 2000;25:2210-2220.
- Reiss S. Expectancy theory of fear, anxiety, and panic. *Clinical Psychology Review* 1991;11:141-153.
- Reneman MF, Jorritsma W, Schellekens JMH, Goeeken LNH. Concurrent validity of questionnaire and performance-based disability measurements in patients with chronic nonspecific low back pain. *Journal of Occupational Rehabilitation* 2002;12:119-129.
- Rief W, Hiller W, Margraf J. Cognitive aspects of hypochondriasis and the somatization syndrome. *Journal of Abnormal Psychology* 1998.
- Rinck M, Becker ES. Approach and avoidance in fear of spiders. *Journal of Behavioral Therapy and Experimental Psychiatry* 2007;38:105-120.
- Rocchi MB, Sisti D, Benedetti P, Valentini M, Bellagamba S, Federici A. Critical comparison of nine different self-administered questionnaires for the evaluation of disability caused by low back pain. *Eura Medicophys* 2005;41:275-281.
- Rode S, Salkovskis PM, Jack T. An experimental study of attention, labelling and memory in people suffering from chronic pain. *Pain* 2001;94:193-203.
- Roelofs J, Goubert L, Vlaeyen JWS, Crombez G. The Tampa Scale for Kinesiophobia: further examination of psychometric properties in patients with chronic low back pain and fibromyalgia. *European Journal of Pain* 2004a;8:495-502.
- Roelofs J, Peters ML, Fassaert T, Vlaeyen JWS. The role of fear of movement and injury in selective attentional processing in patients with chronic low back pain: a dot-probe evaluation. *The Journal of Pain* 2005;6:294-300.
- Roelofs J, Peters ML, Patijn J, Schouten EGW, Vlaeyen JWS. Electronic diary assessment of pain-related fear, attention to pain, and pain intensity in chronic low back pain patients. *Pain* 2004b;112:335-342.
- Roelofs J, Sluiter JK, Frings-Dresen MHW, Goossens MEJB, Thibault P, Boersma K, Vlaeyen JWS. Fear of movement and (re)injury in chronic musculoskeletal pain: evidence for an

- invariant two-factor model of the Tampa Scale for Kinesiophobia across pain diagnoses and Dutch, Swedish, and Canadian samples. *Pain* 2007;131(1-2):181-190.
- Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine* 2000;25:3115-3124.
- Roland M, Morris R. A study of the natural history of back pain. Part I: Development of a reliable and sensitive measure of disability in low back pain. *Spine* 1983;8:141-144.
- Romano JM, Jensen MP, Turner JA, Good AB, Hops H. Chronic pain patient-partner interactions: Further support for a behavioral model of chronic pain. *Behavior Therapy* 2000;31:415-440.
- Rosenstiel AK, Keefe FJ. The use of coping strategies in chronic low back pain patients: Relationship to patient characteristics and current adjustment. *Pain* 1983;17:33-44.
- Rowa K, Antony MM, Swinson RP. Exposure and Response Prevention. In: Antony MM, Purdon C and Summerfeldt LJ (Eds.). *Psychological treatment of obsessive compulsive disorders: Fundamentals and beyond*. Washington, DC: American Psychological Association, 2007. pp. 79-109.
- Rowe MK, Craske MG. Effects of an expanding-spaced vs massed exposure schedule on fear reduction and return of fear. *Behaviour Research and Therapy* 1998a;36:701-717.
- Rowe MK, Craske MG. Effects of varied-stimulus exposure training on fear reduction and return of fear. *Behaviour Research and Therapy* 1998b;36:719-734.
- Rowlands AV, Thomas PW, Eston RG, Topping R. Validation of the RT3 triaxial accelerometer for the assessment of physical activity. *Medicine & Science in Sports & Exercise* 2004;36:518-24.
- Salemink E, van den Hout M, Kindt M. Trained interpretive bias: Validity and effects on anxiety. *Journal of Behavior Therapy and Experimental Psychiatry* 2007;38:212-224.
- Salkovskis PM. Obsessional-compulsive problems: A cognitive-behavioural analysis. *Behaviour Research & Therapy* 1985;23:571-583.
- Sanders SH. operant conditioning with chronic pain: back to basics. In: Turk DC and Gatchel RJ (Eds.). *Psychological approaches to pain management*. New York: The Guilford Press, 2002.
- Schafer JL, Graham JW. Missing data: our view of the state of the art. *Psychological Methods* 2002;7:147-177.
- Schmukle SC, Egloff B. Assessing Anxiety with Extrinsic Simon Tasks. *Experimental Psychology* 2006;53:149-160.
- Schneider C, Palomba D, Flor H. Pavlovian conditioning of muscular responses in chronic pain patients: central and peripheral correlates. *Pain* 2004;112:239-247.
- Schneider W, MEL Professional. In: tools Ps (Ed.), *The Combustion Institute*, Pittsburg, 1996.
- Schoppink LEM, van Tulder MW, Koes BW, Beurskens AJHM, de Bie R. Reliability and validity of the Dutch adaptation of the Quebec Back Pain Disability Scale. *Physical Therapy* 1996;76:268-275.
- Schwartz L, Jensen MP, Romano JM. The development and psychometric evaluation of an instrument to assess spouse responses to pain and well behavior in patients with chronic pain: the Spouse Response Inventory. *The Journal of Pain* 2005;6:243-52.
- Severeijns R, Van den Hout MA, Vlaeyen JWS. The causal status of pain catastrophizing: an experimental test with healthy participants. *European Journal of Pain* 2005;9:257-265.
- Severeijns R, van den Hout MA, Vlaeyen JWS, Picavet HSJ. Pain catastrophising and general health status in a large Dutch community sample. *Pain* 2002;99:367-376.
- Severeijns R, Vlaeyen JWS, Van den Hout MA, Picavet HSJ. Pain catastrophizing is associated with health indices in musculoskeletal pain: a cross-sectional study in the dutch community. *Health Psychology* 2004;23:49-57.
- Severeijns R, Vlaeyen JWS, van den Hout MA, Weber WE. Pain catastrophizing predicts pain intensity, disability, and psychological distress independent of the level of physical impairment. *The Clinical Journal of Pain* 2001;17:165-172.



- Sgroi MI, Willebrand M, Ekselius L, Gerdin B, Andersson G. Fear-avoidance in recovered burn patients: association with psychological and somatic symptoms. *Journal of Health Psychology* 2005;10:491-502.
- Sieben JM, Pain-related fear in acute low back pain: towards understanding and prevention of chronicity (dissertation), Maastricht University, 2005.
- Sieben JM, Portegijs PJM, Vlaeyen JWS, Knottnerus JA. Pain-related fear at the start of a new low back pain episode. *European Journal of Pain* 2005a;9:635-641.
- Sieben JM, Vlaeyen JWS, Portegijs PJM, Verbunt JA, van Riet-Rutgers S, Kester ADM, Von Korff M, Arntz A, Knottnerus JA. A longitudinal study on the predictive validity of the fear-avoidance model in low back pain. *Pain* 2005b;117:162-170.
- Sieben JM, Vlaeyen JWS, Tuerlinckx S, Portegijs PJ. Pain-related fear in acute low back pain: the first two weeks of a new episode. *European Journal of Pain* 2002;6:229-37.
- Silver A, Haeny M, Vijayadurai P, Wilks D, Pattrick M, Main CJ. The role of fear of physical movement and activity in chronic fatigue syndrome. *Journal of Psychosomatic Research* 2002;52:485-493.
- Smeets RJ, Hijdra HJ, Kester AD, Hitters MW, Knottnerus JA. The usability of six physical performance tasks in a rehabilitation population with chronic low back pain. *Clinical Rehabilitation* 2006a;20:989-97.
- Smeets RJ, Vlaeyen JW, Hidding A, Kester AD, van der Heijden GJ, van Geel AC, Knottnerus JA. Active rehabilitation for chronic low back pain: Cognitive-behavioral, physical, or both? First direct post-treatment results from a randomized controlled trial [ISRCTN22714229]. *BMC Musculoskeletal Disorders* 2006b;7:5.
- Smeets RJ, Wittink H, Hidding A, Knottnerus JA. Do patients with chronic low back pain have a lower level of aerobic fitness than healthy controls?: Are pain, disability, fear of injury, working status, or level of leisure time activity associated with the difference in aerobic fitness level? *Spine* 2006c;31:90-98.
- Smeets RJEM, Beelen S, Goossens MEJB, Schouten EGW, Knottnerus JA, Vlaeyen JWS. Treatment expectancy and credibility are associated with the outcome of both physical and cognitive-behavioral treatment in chronic low back pain. *The Clinical Journal of Pain*, In press.
- Smeets RJEM, Vlaeyen JWS, Hidding A, Kester ADM, van der Heijden GJMG, Knottnerus JA. Chronic low back pain: Physical training, graded activity with problem solving training, or both? The one-year post-treatment results of a randomized controlled trial. *Pain* 2007 (doi:10.1016/j.pain.2007.04.021).
- Smeets RJEM, Vlaeyen JWS, Kester ADM, Knottnerus JA. Reduction of Pain Catastrophizing Mediates the Outcome of Both Physical and Cognitive-Behavioral Treatment in Chronic Low Back Pain. *The Journal of Pain* 2006d;7:261-271.
- Sorbi MJ, Peters ML, Kruise DA, Maas CJ, Kerssens JJ, Verhaak PF, Bensing JM. Electronic momentary assessment in chronic pain II: psychological pain responses as predictors of pain disability. *The Clinical Journal of Pain* 2006;22:67-81.
- Spenkeliink CD, Hutten MM, Hermens HJ, Greitemann BO. Assessment of activities of daily living with an ambulatory monitoring system: a comparative study in patients with chronic low back pain and nonsymptomatic controls. *Clinical Rehabilitation* 2002;16:16-26.
- Spinhoven P, ter Kuile M, Kole-Snijders AMJ, Hutten Mansfeld M, den Ouden DJ, J.W.S. V. Catastrophizing and internal pain control as mediators of outcome in the multidisciplinary treatment of chronic low back pain. *European Journal of Pain* 2004;8:211-219.
- SPSS-Inc, SPSS for Windows (Version 11.0.1), SPSS Inc, Chicago, 2001.
- Staerkle R, Mannion AF, Elfering A, Junge A, Semmer NK, Jacobshagen N, Grob D, Dvorak J, Boos N. Longitudinal validation of the fear-avoidance beliefs questionnaire (FABQ) in

- a Swiss-German sample of low back pain patients. *European Spine Journal* 2004;13:332-340.
- Stayhealthy, RT3 Tri-axial Research Tracker, Stayhealthy Inc., Elkader, IA.
- Sterling M, Kenardy J, Jull G, Vicenzino B. The development of psychological changes following whiplash injury. *Pain* 2003;106:481-489.
- Storheim K, Brox JI, Holm I, Bo K. Predictors of return to work in patients sick listed for sub-acute low back pain: a 12-month follow-up study. *Journal of Rehabilitation Medicine* 2005;37.
- Stratford PW, Binkley JM, Riddle DL, Guyatt GH. Sensitivity to change of the Roland-Morris Back Pain Questionnaire: Part 1. *Physical Therapy* 1998;78:1186-1196.
- Streiner DL. Finding our way: an introduction to path analysis. *Canadian Journal of Psychiatry; Revue canadienne de psychiatrie* 2005;50:115-22.
- Sullivan MJ, Neish N. The effects of disclosure on pain during dental hygiene treatment: the moderating role of catastrophizing. *Pain* 1999;79:155-163.
- Sullivan MJ, Rodgers WM, Wilson PM, Bell GJ, Murray TC, Fraser SN. An experimental investigation of the relation between catastrophizing and activity intolerance. *Pain* 2002;100:47-53.
- Sullivan MJ, Stanish WD. Psychologically based occupational rehabilitation: the Pain-Disability Prevention Program. *The Clinical Journal of Pain* 2003;19:97-104.
- Sullivan MJL, Bishop SR, Pivik J. The Pain Catastrophizing Scale: Development and validation. *Psychological Assessment* 1995;7:524-532.
- Sullivan MJL, Lynch ME, Clark AJ. Dimensions of catastrophic thinking associated with pain experience and disability in patients with neuropathic pain conditions. *Pain* 2005a;113:310-315.
- Sullivan MJL, Thorn B, Haythornthwaite JA, Keefe F, Martin M, Bradley LA, Lefebvre JC. Theoretical perspectives on the relation between catastrophizing and pain. *The Clinical Journal of Pain* 2001;17:52-64.
- Sullivan MJL, Thorn B, Rodgers W, Ward LC. Path model of psychological antecedents to pain experience: experimental and clinical findings. *The Clinical Journal of Pain* 2004;20:164-73.
- Sullivan MJL, Ward LC, Tripp D, French DJ, Adams H, Stanish WD. Secondary prevention of work disability: community-based psychosocial intervention for musculoskeletal disorders. *Journal of Occupational Rehabilitation* 2005b;15:377-392.
- Swinkels-Meewisse EJCM, Swinkels RAHM, Verbeek ALM, Vlaeyen JWS, Oostendorp RAB. Psychometric properties of the Tampa Scale for kinesiophobia and the fear-avoidance beliefs questionnaire in acute low back pain. *Manual Therapy* 2003a;8:29-36.
- Swinkels-Meewisse IE, Roelofs J, Oostendorp RA, Verbeek AL, Vlaeyen JWS. Acute low back pain: pain-related fear and pain catastrophizing influence physical performance and perceived disability. *Pain* 2006a;120:36-43.
- Swinkels-Meewisse IE, Roelofs J, Verbeek AL, Oostendorp RA, Vlaeyen JWS. Fear of movement/(re)injury, disability and participation in acute low back pain. *Pain* 2003b;105:371-379.
- Swinkels-Meewisse IE, Roelofs J, Verbeek AL, Oostendorp RA, Vlaeyen JWS. Fear-avoidance beliefs, disability, and participation in workers and non-workers with acute low back pain. *The Clinical Journal of Pain* 2006b;22:45-54.
- Swinkels-Meewisse IEJ, Roelofs J, Schouten EGW, Verbeek ALM, Oostendorp RAB, Vlaeyen JWS. Fear of movement/(re)injury predicting chronic disabling low back pain: a prospective inception cohort study. *Spine* 2006c;31:658-664.
- Symonds TL, Burton AK, Tilotson KM, Main CJ. Absence resulting from low back trouble can be reduced by psychosocial intervention at the work place. *Spine* 1995;20:2738-2745.
- Teachman BA. Evaluating implicit spider fear associations using the Go/No-go Association Task. *Journal of Behavior Therapy and Experimental Psychiatry* 2007;38:156-167.

- Teige S, Schnabel K, Banse R, Asendorpf JB. Assessment of multiple implicit self-concept dimensions using the Extrinsic Affective Simon Task (EAST). *European Journal of Personality* 2004;18:495-520.
- Thorn BE, Cross TH, Walker BB. Meta-analyses and systematic reviews of psychological treatments for chronic pain: relevance to an evidence-based practice. *Health psychology official journal of the Division of Health Psychology, American Psychological Association* 2007;26:10-2.
- Thush C, Wiers RW. Explicit and implicit alcohol-related cognitions and the prediction of future drinking in adolescents. *Addictive Behaviors* 2007;32:1367-1383.
- Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. *The Journal of the American Medical Association* 2003;290:1624-32.
- Turk DC. The Potential of Treatment Matching for Subgroups of Patients With Chronic Pain: Lumping Versus Splitting. *The Clinical Journal of Pain* 2005;21:44-55.
- Turk DC, Robinson JP, Burwinkle T. Prevalence of fear of pain and activity in patients with fibromyalgia syndrome. *The Journal of Pain* 2004;5:483-490.
- Turner JA, Franklin G, Fulton Kehoe D, Sheppard L, Wickizer TM, Wu R, Gluck JV, Egan K. Worker recovery expectations and fear-avoidance predict work disability in a population-based workers' compensation back pain sample. *Spine* 2006;31:682-9.
- Turner JA, Holtzman S, Mancle L. Mediators, moderators, and predictors of therapeutic change in cognitive-behavioral therapy for chronic pain. *Pain* 2007;127.
- Turner JA, Jensen MP, Romano JM. Do beliefs, coping, and catastrophizing independently predict functioning in patients with chronic pain? *Pain* 2000;85:115-125.
- Turner JA, Jensen MP, Warms CA, Cardenas DD. Catastrophizing is associated with pain intensity, psychological distress, and pain-related disability among individuals with chronic pain after spinal cord injury. *Pain* 2002;98:127-34.
- Turner JA, Mancl L, Aaron LA. Pain-related catastrophizing: a daily process study. *Pain* 2004;103-111.
- van Breukelen GJP. ANCOVA vs. change from baseline: more power in randomized studies, more bias in nonrandomized studies. *Journal of Clinical Epidemiology* 2006;59:920-925.
- Van Breukelen GJP, Vlaeyen JWS. Norming clinical questionnaires with multiple regression: The Pain Cognition List. *Psychological Assessment* 2005;17:336-344.
- Van Damme S, Crombez G, Bijttebier P, Goubert L, Van Houdenhove B. A confirmatory factor analysis of the Pain Catastrophizing Scale: invariant factor structure across clinical and non-clinical populations. *Pain* 2002a;96:319-324.
- Van Damme S, Crombez G, Eccleston C. Retarded disengagement from pain cues: the effects of pain catastrophizing and pain expectancy. *Pain* 2002b;100:111-118.
- Van Damme S, Crombez G, Eccleston C. The anticipation of pain modulates spatial attention: evidence for pain-specificity in high-pain catastrophizers. *Pain* 2004a;111:392-399.
- Van Damme S, Crombez G, Eccleston C. Disengagement from pain: the role of catastrophic thinking about pain. *Pain* 2004b;107:70-76.
- Van Damme S, Crombez G, Hermans D, Koster EH, Eccleston C. The role of extinction and reinstatement in attentional bias to threat: A conditioning approach. *Behaviour Research and Therapy* 2006;44:1555-1563.
- Van Damme S, Crombez G, Vlaeyen JWS, Goubert L, Van den Broeck A, Van Houdenhove B. De Pain Catastrophizing Scale: Psychometrische karakteristieken en normering. / The Pain Catastrophizing Scale: Psychometric characteristics and norms. *Gedragstherapie* 2000;33:209-220.
- van den Berg-Emons RJ, Schasfoort FC, de Vos LA, Bussmann JB, Stam HJ. Impact of chronic pain on everyday physical activity. *European Journal of Pain* 2007;11:587-593.

- van den Hout A, To solve or not to solve? Effects of problem solving therapy and graded activity in non-specific low back pain, University of Maastricht, Maastricht, 2002.
- van den Hout JH, Vlaeyen JW, Heuts PH, Zijlema JH, Wijnen JA. Secondary prevention of work-related disability in nonspecific low back pain: does problem-solving therapy help? A randomized clinical trial. *The Clinical Journal of Pain* 2003;19:87-96.
- van den Hout JHC, Vlaeyen JWS, Heuts PHTG, Sillen WJH, Willen AJEHL. Functional disability in non-specific low back pain: the role of pain-related fear and problem-solving skills. *International Journal of Behavioral Medicine* 2001a;8:149-162.
- van den Hout JHC, Vlaeyen JWS, Houben RMA, Soeters APM, Peters ML. The effects of failure feedback and pain-related fear on pain report, pain tolerance and pain avoidance in chronic low back pain patients. *Pain* 2001b;92:247-257.
- van der Zee KI, Sanderman R, Het meten van gezondheidstoestand met de RAND-36. Een handleiding, Groningen, 1993.
- van der Zee KI, Sanderman R, Heyink JW. De psychometrische kwaliteiten van de MOS 36-item Short Form Health Survey(SF-36) in een Nederlandse populatie. *Tijdschrift Sociale Gezondheidszorg* 1993;71:183-191.
- van der Zee KI, Sanderman R, Heyink JW, de Haes H. Psychometric qualities of the RAND-36-item health survey 1.0: a multidimensional measure of general health status. *International Journal of Behavioral Medicine* 1996;3:104-122.
- Van Nieuwenhuysse A, Somville PR, Crombez G, Burdorf A, Verbeke G, Johannik K, Van den Bergh O, Masschelein R, Mairiaux P, Moens GF. The role of physical workload and pain related fear in the development of low back pain in young workers: evidence from the BelCoBack Study; results after one year of follow up. *Occupational and Environmental Medicine* 2006;63:45-52.
- van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in The Netherlands. *Pain* 1995;62:233-40.
- van Tulder MW, Koes BW, Bouter LM. Conservative treatment of acute and chronic nonspecific low back pain: a systematic review of randomized controlled trials of the most common interventions. *Spine* 1997;22:2128-2156.
- van Tulder MW, Ostelo R, Vlaeyen JW, Linton SJ, Morley SJ, Assendelft WJ. Behavioral treatment for chronic low back pain: a systematic review within the framework of the Cochrane Back Review Group. *Spine* 2001;26:270-81.
- Vancleef LM, Peters ML, Gilissen SM, De Jong PJ. Understanding the role of injury/illness sensitivity and anxiety sensitivity in (automatic) pain processing: an examination using the extrinsic affective Simon task. *The Journal of Pain* 2007;8:563-72.
- Vancleef LMG, Peters ML. Pain catastrophizing, but not injury/illness sensitivity or anxiety sensitivity, enhances attentional interference by pain. *The Journal of Pain* 2006;7:23-30.
- Vancleef LMG, Peters ML, Roelofs J, Asmundson GJG. Do fundamental fears differentially contribute to pain-related fear and pain catastrophizing? An evaluation of the sensitivity index. *European Journal of Pain* 2006;10:527-536.
- Verbeke G, Molenberghs G, Linear mixed models for longitudinal data, Springer-Verlag, New York, 2000.
- Verbunt JA, Seelen HAM, Vlaeyen JWS, Bousema EJ, van der Heijden GJ, Heuts PHTG, Knottnerus JA. Pain-related factors contributing to muscle inhibition in patients with chronic low back pain: an experimental investigation based on superimposed electrical stimulation. *The Clinical Journal of Pain* 2005;21:232-240.
- Verbunt JA, Seelen HAM, Vlaeyen JWS, van der Heijden GJ, Heuts PHTG, Pons K, Knottnerus JA. Disuse and deconditioning in chronic low back pain: concepts and hypotheses on contributing mechanisms. *European Journal of Pain* 2003a;7:9-21.
- Verbunt JA, Seelen HAM, Vlaeyen JWS, van der Heijden GJ, Knottnerus JA. Fear of injury and physical deconditioning in patients with chronic low back pain. *Archives Of Physical Medicine And Rehabilitation* 2003b;84:1227-1232.

- Verbunt JA, Westerterp KR, van der Heijden GJ, Seelen HA, Vlaeyen JWS, Knottnerus JA. Physical activity in daily life in patients with chronic low back pain. *Archives Of Physical Medicine And Rehabilitation* 2001;82:726-730.
- Vermilyea BB, Barlow DH, O'Brien GT. The importance of assessing treatment integrity: An example in the anxiety disorders. *Journal of Psychopathology and Behavioral Assessment* 1984.
- Vlaeyen JW, De Jong JR, Onghena P, Kerckhoffs-Hanssen M, Kole-Snijders AM. Can pain-related fear be reduced? The application of cognitive- behavioural exposure in vivo. *Pain Research and Management* 2002a;7:144-153.
- Vlaeyen JW, Geurts SM, Kole-Snijders AM, Schuerman JA, Groenman NH, van Eek H. What do chronic pain patients think of their pain? Towards a pain cognition questionnaire. *British Journal of Clinical Psychology* 1990;29 ( Pt 4):383-94.
- Vlaeyen JWS, Crombez G. The modified Tampa Scale for Kinseophobia: for its use in non-pain samples. Unpublished Dutch/Flemish translation. 1998.
- Vlaeyen JWS, de Jong J, Geilen M, Heuts PHTG, van Breukelen G. Graded exposure in vivo in the treatment of pain-related fear: a replicated single-case experimental design in four patients with chronic low back pain. *Behaviour Research and Therapy* 2001;39:151-166.
- Vlaeyen JWS, de Jong J, Geilen M, Heuts PHTG, van Breukelen G. The treatment of fear of movement/(re)injury in chronic low back pain: Further evidence on the effectiveness of exposure in vivo. *The Clinical Journal of Pain* 2002b;18:251-261.
- Vlaeyen JWS, de Jong J, Leeuw M, Crombez G. Fear reduction in chronic pain: graded exposure in vivo with behavioral experiments. In: Asmundson GJ, Vlaeyen JWS and Crombez G (Eds.). *Understanding and Treating Fear of Pain*. New York: Oxford University Press, 2004a. pp. 313-343.
- Vlaeyen JWS, de Jong J, Sieben JM, Crombez G. Graded exposure in vivo for pain-related fear. In: Turk DC and Gatchel RJ (Eds.). *Psychological Approaches to Pain Management. A practitioner's handbook*. New York: The Guilford Press, 2002c.
- Vlaeyen JWS, Kole Snijders AMJ, Boeren RGB, van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain* 1995a;62:363-372.
- Vlaeyen JWS, Kole Snijders AMJ, Rotteveel AM, Ruesink R, Heuts PHTG. The role of fear of movement/(re)injury in pain disability. *Journal of Occupational Rehabilitation* 1995b;5:235-252.
- Vlaeyen JWS, Kole-Snijders AMJ, Crombez G, Boeren R, Rotteveel AM. De Tampa Schaal voor Kinesiofobie TSK, Nederlands geautoriseerde versie. 1995c.
- Vlaeyen JWS, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. *Pain* 2000;85:317-332.
- Vlaeyen JWS, Morley S. Active despite pain: the putative role of stop-rules and current mood. *Pain* 2004;110:512-516.
- Vlaeyen JWS, Morley S. Cognitive-behavioral treatments for chronic pain: what works for whom? *The Clinical Journal of Pain* 2005;21:1-8.
- Vlaeyen JWS, Timmermans C, Rodriguez LM, Crombez G, van Horne W, Ayers GM, Albert A, Wellens HJ. Catastrophic thinking about pain increases discomfort during internal atrial cardioversion. *Journal of Psychosomatic Research* 2004b;56:139-144.
- von Korff M, Dworkin SF, Le Resche LA, Kruger A. An epidemiologic comparison of pain complaints. *Pain* 1988;32:173-183.
- Vowles KE, Gross RT. Work-related beliefs about injury and physical capability for work in individuals with chronic pain. *Pain* 2003;101:291-298.
- Waddell G. Low back pain: a twentieth century health care enigma. *Spine* 1996;21:2820-2825.
- Waddell G, *The Back Pain Revolution*, Churchill Livingstone, 1998a.

- Waddell G. Diagnostic triage. In: Waddell G (Eds.). *The back pain revolution*. Edinburgh: Churchill Livingston, 1998b. pp. 9-25.
- Waddell G. The biopsychosocial model. In: Waddell G (Eds.). *The Back Pain Revolution*. Edinburgh: Churchill Livingston, 2004a. pp. 265-282.
- Waddell G. Diagnostic triage. In: Waddell G (Eds.). *The Back Pain Revolution*. Edinburgh: Churchill Livingston, 2004b. pp. 9-26.
- Waddell G. The epidemiology of back pain. In: Waddell G (Eds.). *The Back Pain Revolution*. Edinburgh: Churchill Livingston, 2004c. pp. 71-89.
- Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain* 1993;52:157-168.
- Walker BF. The prevalence of low back pain: a systematic review of the literature from 1966 to 1998. *Journal of Spinal Disorders* 2000;13:205-217.
- Wall PD. On the relation of injury to pain. The John J. Bonica lecture. *Pain* 1979;6:253-264.
- Waltz J, Addis ME, Koerner K, Jacobson NS. Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology* 1993;61:620-630.
- Wampold BE. *The Great Psychotherapy Debate: Models, Methods, and Findings*, Lawrence Erlbaum Associates, London, 2000.
- White B, Sanders SH. The influence on patients' pain intensity ratings of antecedent reinforcement of pain talk or well talk. *Journal of Behavior Therapy and Experimental Psychiatry* 1986;17:155-159.
- Wittink H, Hoskins Michel T, Wagner A, Sukiennik A, Rogers W. Deconditioning in patients with chronic low back pain: fact or fiction? *Spine* 2000;25:2221-2228.
- Wittink H, Michel TH, Sukiennik A, Gascon C, Rogers W. The association of pain with aerobic fitness in patients with chronic low back pain. *Archives Of Physical Medicine And Rehabilitation* 2002;83:1467-1471.
- Woby SR, Watson PJ, Roach NK, Urmston M. Adjustment to chronic low back pain - the relative influence of fear-avoidance beliefs, catastrophizing, and appraisals of control. *Behaviour Research and Therapy* 2004a;42:761-774.
- Woby SR, Watson PJ, Roach NK, Urmston M. Are changes in fear-avoidance beliefs, catastrophizing, and appraisals of control, predictive of changes in chronic low back pain and disability? *European Journal of Pain* 2004b;8:201-210.
- Woods MP, Asmundson GJG. Evaluating the efficacy of graded in vivo exposure for the treatment of fear in patients with chronic back pain: A randomized controlled clinical trial. *Pain* 2007 (doi:10.1016/j.pain.2007.06.037 ).
- Yeaton WH, Sechrest L. Critical dimensions in the choice and maintenance of successful treatments: Strength, integrity, and effectiveness. *Journal of Consulting and Clinical Psychology* 1981;49:156-167.
- Zvolensky MJ, Goodie JL, McNeil DW, Sperry JA, Sorrell JT. Anxiety sensitivity in the prediction of pain-related fear and anxiety in a heterogeneous chronic pain population. *Behaviour Research and Therapy* 2001;39:683-696.



## **Dankwoord**





Het boekje is bijna uit. Ik ben aangekomen bij het hoofdstuk dat niet alleen over het algemeen het meest gelezen wordt, maar wat voor mij ook een zeer waardevol onderdeel van dit proefschrift is gebleken. De voltooiing van dit boekje markeert voor mij namelijk het einde van een bijzondere, maar soms ook moeizame periode, waarin ik niet alleen veel over chronische pijn en wetenschappelijk onderzoek, maar ook over mezelf, heb geleerd. Maar het belangrijkste van dit hoofdstuk is dat ik hierin uitgebreid de ruimte kan nemen om alle lieve mensen te bedanken die direct of indirect hebben bijgedragen aan de totstandkoming van dit proefschrift.

Als pijnonderzoeker onder de vleugels van Johan Vlaeyen heb ik de luxe ervaren om mijn onderzoek uit te voeren binnen een hecht netwerk van andere pijnonderzoekers. Hierdoor zijn er vele mensen inhoudelijk betrokken geweest bij de uitvoering van mijn onderzoek.

Allereerst wil ik natuurlijk mijn dagelijks begeleider en promotor Johan Vlaeyen, en mijn andere dagelijks begeleider en copromotor Mariëlle Goossens, bedanken. Ik begin met enkele woorden van dank die op jullie beiden van toepassing zijn. Tijdens onze regelmatige overleggen hebben wij gedrietjes mijn promotieonderzoek vorm gegeven. Jullie dachten intensief mee over de uitvoering van de projecten en over oplossingen voor de oneindige stroom van problemen die onlosmakelijk verbonden zijn met een multicenter RCT. Daarnaast stimuleerden jullie me om naast deze hoofdstudie andere studies op te zetten en uit te voeren, waardoor ik aan verschillende soorten onderzoek heb kunnen proeven. Ik heb dankzij jullie geleerd een kritisch en gedegen onderzoeker te zijn. Jullie waren er wanneer ik jullie nodig had, maar lieten me vooral mijn gang gaan wanneer dat niet het geval was. Daarbij hebben jullie me altijd veel ruimte en vrijheid geboden. Daarnaast heb ik jullie aandacht voor persoonlijke zaken altijd zeer waardevol gevonden. Jullie respect voor mijn eigen inbreng en waardering voor mij als persoon hebben me veel vertrouwen gegeven. Dank jullie wel dat ik altijd onder prettige omstandigheden heb kunnen werken.

Dan enkele woorden voor jullie apart. Johan, jij hebt ondanks je drukke agenda altijd tijd voor me vrijgemaakt om me te begeleiden. Vooral op het einde van het traject was ik blij verrast dat je mijn eindsprint wist bij te benen en dat ik binnen enkele dagen de teksten voorzien van commentaar terugkreeg. Ik bewonder je passie en inzet voor het onderzoek en je rustige, maar stimulerende en prikkelende begeleiding van je promovendi. Ongeacht de omvang van de drukte, straal je een rust uit die aanstekelijk is op anderen. Dankzij jouw roze bril, je rooskleurige visie op eigenlijk alles, heb ik geleerd om zaken meer te relativiseren en van een positieve kant te bekijken. Dit was een zeer prettige, maar misschien ook noodzakelijke, bijkomstigheid in de uitvoering van een promotietraject. Mariëlle, jij bent bijgesprongen toen Johan's takenpakket bleef groeien waardoor er mogelijk onvoldoende tijd voor mijn begeleiding zou resteren. Het was erg fijn dat jouw deur altijd openstond en dat er gelegenheid was tot begeleiding, maar vaak ook voor gewoon een gezellig praatje op momenten die daar om vroegen.

Sita, zonder jou was het BackActive project nooit zo goed gelopen als het heeft gedaan. Je bent de koningin van de duizendpoten en het was een heerlijke zekerheid dat jij zorg droeg voor de dagelijkse gang van zaken van het project, vooral vanwege het vertrouwen dat ik kon hebben in jouw nauwkeurigheid. Je dacht mee op bijna alle

vlakken van mijn promotieonderzoek. Niet alleen regelde je jouw eigen bezigheden, maar regelmatig hielp je mij ook nog aan belangrijke zaken herinneren. Vooral doordat je dit altijd hebt weten te combineren met een enorme dosis heerlijke humor en plezier is samenwerken met jou altijd geweldig geweest. Daarbij hebben we ook altijd heerlijk een eind weg kunnen bomen over andere dingen dan werk (begin maar niet over de kleinkinderen!). Oneindig veel dank dus voor de vele bezigheden die je voor het project en mij hebt willen uitvoeren, maar vooral voor je vrolijke lach, je gezelligheid en je moederlijke advies op de momenten die daar om vroegen.

Saskia, ook jij hebt een belangrijke bijdrage aan BackActive geleverd. Niet alleen heb je zonder gemopper als ‘vliegende keep’ gefungeerd en ben je met het openbaar vervoer (alleen dat al is een complimentje waard, vooral naar Hoensbroek!) als psycholoog ingesprongen in drie behandelcentra, maar ook heb je daarnaast de nodige onderzoeksassistentie verleend. Je hebt hiervoor tienduizenden VAS-schaaltjes gescoord en je bent op vele vlakken nauw betrokken geweest bij de uitvoering van het project. Een groot dank je wel voor al je verzette werk (en daarnaast natuurlijk ook voor je vriendschap, maar daar kom ik later nog op terug).

Madelon, Jeroen, en Peter, jullie waren lid van de BackActive projectgroep en zijn tijdens het opstarten en looptijd van het BackActive project een waardevolle aanvulling geweest op de dagelijkse begeleiding. Jullie hebben de voortgang van het project altijd nauwlettend gevolgd en hebben aan alle belangrijke beslissingen bijgedragen die genomen moesten worden. Bedankt voor jullie tijd en kritische inbreng

Bij de daadwerkelijk uitvoering van het BackActive project in de verschillende behandelcentra zijn ook veel mensen betrokken geweest.

De behandelingen in het kader van het BackActive project zijn uitgevoerd in samenwerking met vier behandelcentra, namelijk de afdeling revalidatie Academisch Ziekenhuis Maastricht, het revalidatiecentrum Hoensbroeck, de afdeling revalidatie Maaslandziekenhuis Sittard, en het revalidatiecentrum Blixembosch. Vanuit deze centra, maar ook vanuit de pijnpoli van het Academisch Ziekenhuis Maastricht, de afdeling revalidatie van het Laurentiusziekenhuis Roermond, de afdeling revalidatie van het Atriumziekenhuis Heerlen, en de Arbodienst Limburg, hebben vele artsen zorggedragen voor patiëntaanmeldingen voor het project. Daarnaast zijn er in deze centra veel behandelaars en logistiek ondersteuners geweest die zich, ondanks vele hobbels en tegenslagen, met enthousiasme zijn blijven inzetten voor het BackActive project. Ik wil graag iedereen hiervoor bedanken (met het risico om iemand te vergeten te noemen, waarvoor alvast excuses): *De artsen:* Joop Ruijgrok, Jaap Patijn, Marion van Eijssen, Peter Heuts, Martin Danen, Anita Tinga, Geerie Winnubst, Rob Smeets, Peter van Leeuwen, Jeanine Verbunt, Edwin Drossaer, Erik de Klerk, Monique Steenis, Slavicha Basic. *De psychologen, fysiotherapeuten en ergotherapeuten:* Saskia Nijst, Jeroen de Jong, Ruud Severeijns, Nienke Verkerk, Ank Kole, Jerome Korse, Theo Verbunt, Anja van den Hout, Marlies den Hollander, Christoph Loo, Corine Cuijpers, Anita Stevens, Mario Geilen, Herman Mulder, Carla Hoogervorst, Carla Palmen, Kim Kugler, Iris Kanera, Rien Berrevoets, Mark Baudoin. *De logistiek ondersteuners:* Els Oostveen, Jannique Geurten, Coby Scheijen, Peter Kastrop, Hans Veders, Roger Bastiaens, Magda van Duinen, Inge Lamé, Elke Vermeulen, Iris van Koetsveld.

De trainingen en interviews van beide behandelingen zijn verzorgd door Jeroen en Albère. Albère heeft daarnaast een belangrijke inzet geleverd door de instroom van patiënten te optimaliseren. Ook Jeroen heeft altijd veel voor de lopende zaken van BackActive betekend. Allebei bedankt hiervoor!

Ook dank aan Magda, omdat je in Blixembosch als een soort van “Sita Twee” de uitvoering van de metingen bij de BackActive patiënten aldaar op je hebt genomen.

Natuurlijk zou dit proefschrift er nooit zijn gekomen wanneer niet vele patiënten bereid zouden zijn geweest om deel te nemen aan de metingen en de behandelingen. Mijn grote dank gaat dus ook uit naar deze anonieme groep van mensen die de moeite hebben genomen om meerdere keren naar de testlocaties te komen en thuis de dagboeken in te vullen en de activiteitenmeter te dragen.

De laatste groep mensen die beroepsmatig hebben bijgedragen aan mijn onderzoek hebben dat op diverse manieren gedaan. Gerard, zonder jouw hoogdravende, maar ontzettend heldere, uitleg zou ik de statische analyses van twee studies niet hebben kunnen uitvoeren. Jij weet statistiek boeiend en uitdagend te maken en ik vond onze samenwerking erg leerzaam. Ik ben je reuze dankbaar voor de vele uren die je aan mijn statistische begeleiding hebt besteed en ik bewonder de snelheid waarmee je mijn honderden vragen per mail hebt beantwoord. Ook een groot woord van dank aan Rosanne, dankzij wie de vragenlijsten elektronisch afgenomen hebben kunnen worden. Hierdoor is mij erg veel werk bespaard gebleven. Charlie, jij bedankt voor het programmeren van de EAST en de GNAT voor de impliciete studie. Jan Jochijms en Jolanda Idink van Top-Care Maastricht Airport, dank jullie wel voor het aanleveren van patiënten en het beschikbaar stellen van jullie apparatuur voor de impliciete studie. Daarnaast een dank je wel aan Eric, Henk, Jeanine en Ivan voor het meedenken over de analyses van de activiteitengegevens. Dankzij jullie was deze nachtmerrie nog enigszins te hanteren. Susan Picavet, dank je wel voor het beschikbaar stellen van de dataset van de DMC<sub>3</sub> studie, waardoor ik in staat ben geweest om in een vroeg stadium van mijn promotie alvast een artikel te schrijven. Erik, jij bedankt voor je eeuwige beschikbaarheid voor vragen over statistiek. Alle coauteurs, bedankt voor jullie constructieve bijdrage aan de artikelen. De SOMATO leden, jullie bedankt voor de leerzame en inspirerende tweewekelijkse bijeenkomsten. Dank je wel Astrid, voor het controleren van de scoring en invoer van de vele dagelijkse dagboekjes. Angela en iedereen van het secretariaat, bedankt voor alle regelingen achter de schermen. Sarah en Dominique, dank jullie wel voor jullie betrouwbare en snelle inzet in het beluisteren en scoren van de bandjes voor de behandelintegriteit. Floske en Lianne, jullie bedankt voor de dataverzameling en verwerking tijdens jullie onderzoeksstages.

Dan een tussenstapje tussen de mensen die ik bedanken wil vanuit hun beroepsmatige bijdrage aan mijn proefschrift en die uit mijn privé-leven, namelijk de grote groep gezellige collega's waardoor ik het altijd erg naar mijn zin heb gehad op de afdeling DMKEP/CPS.

Ruud, je was bijna 3 jaar mijn kamergenoot en je zult wel gek zijn geworden van al mijn vragen en geklets. We hebben een gezellige tijd gehad, waarbij je voor mij als een voorbeeld bent geweest, niet alleen omdat je alles (alles? Ja, alles!) weet, maar vooral omdat je een heel rustgevende en relativerend effect op me hebt weten te hebben. Ik vind het fijn dat we nog altijd contact hebben. Johanna, jij bent het laatste

jaar mijn kamergenoot geweest en ook dat was altijd zeer gezellig. We hebben samen heel wat (klets)pauses genomen en Douwe Egberts koffiekaarten vol gestempeld. Jeffrey, bij jou mag ik altijd binnenlopen en je lastigvallen voor een praatje of een domme vraag, die weliswaar niet bestaan, maar die ik vooral aan jou durf te stellen. Lottuuuuuuuuuuuuuh, jij ook bedankt voor de vele babbels, mailtjes en wandelgangetjes naar de keuken voor een potje thee. Natuurlijk ook alle promovendi bedankt voor de gezelligheid, de leuke EPP dagen, de inhoudelijke overleggen en de lekkere AIO-etentjes. En niet te vergeten de tafeltennissers, die voor een welkome afleiding zorgden wanneer ik in een afternoon-dipje zat, maar ook de 'lunchers' die zich altijd stipt om 13.00 richting AZM restaurant begaven. Ook heb ik samen met de Wijze Wielren Wijven club meerdere zomers het geweldige Limburgse landschap leren ontdekken. Ik hoop jullie volgende zomer weer te zien!

Dit proefschrift is natuurlijk niet alleen tot stand gekomen dankzij de vele mensen van de werkvloer, maar ook door de bijdragen van geweldige familieleden en vrienden.

Karian en Christianne, 'de mier' en 'de eekhoorn', jullie vriendschap is van onschatbare waarde. Ik, 'de olifant', vind het heerlijk om aan de lamp te slingeren tijdens de vele etentjes in het dierenbos, waarbij we samen het leven overpeinzen, maar bovenal veel lol hebben en (om elkaar) lachen. Het is een fijn gevoel dat we er voor elkaar zijn en meelevens met elkaar's wel en wee. Ik bewonder dat jullie kleine beestjes me altijd weten op te vangen wanneer ik weer eens uit de boom val (en het is helemaal een wonder dat een mier en eekhoorn geen rugklachten ontwikkelen na het opvangen van een olifant). Des te bijzonderder vind ik het dat jullie als paranimfen naast me willen staan tijdens mijn promotie. Ik kan me geen mooiere dag en leuker feest wensen dan met jullie tweetjes aan mijn zij.

Afke, ook al zien we elkaar niet zoveel meer door de afstand tussen Utrecht en Maastricht, toch is het contact hierdoor eerder intenser geworden dan minder. We houden elkaar op de hoogte van alles wat ons bezighoudt, maar ook nog veel meer dan dat, door vele mailtjes en telefoontjes. Ik kan veel leren van jouw kijk op dingen die soms zo tegengesteld is aan de mijne. Ik bewonder je eeuwige lach en positieve, maar ook nuchtere, instelling. Ik vind het geweldig dat ik onvoorwaardelijk op je kan rekenen en dat ik met werkelijk alles bij je terecht kan.

Saskia, dank je wel voor onze vriendschap, waarin we een hoop lol beleven aan reflectie op onze eigen catastroferende cognities. Ook weet ik mezelf mede dankzij jou twee keer per week richting sportschool te begeven. Ik hoop onze gezellige etentjes voort te kunnen nu we niet langer meer UM collega's zijn. Gezellig dat we in Hoensbroek alsnog collega's blijven. Femmie en Marlies, hoewel we onze AAC etentjes met M&M's en wijn hebben moeten inruilen voor een wat minder frequent contact, vind ik het leuk dat we elkaar nog steeds regelmatig zien! Bas en Monique, oud-huisgenoten van 'Kasteel de Daelende Broeck', leuk dat er elkaar altijd nog zijn blijven spreken en zien!

Ten aanzien van mijn familie wil ik als eerste graag mijn ouders bedanken, aangezien ik dankzij hen ben wie ik ben. Pap en mam, jullie hebben me altijd gesteund, en wat ik ook deed, jullie zijn trots op me geweest. Jullie hebben me meegegeven dat het niet de prestatie is, maar de intentie en inzet die tellen, en dat ik vooral bij mezelf moet blijven. Ik ben goed om wie ik ben en dat geeft een heel bijzonder gevoel! Ik mag altijd even bellen als mijn mentale emmertje overstroomt en

dan staan jullie klaar met een zachte schouder en wijs advies. Dankzij jullie heb ik geleerd om bij tegenslag door te zetten en niet op te geven. Zonder deze eigenschap vraag ik me af of dit proefschrift hier zou hebben gelegen.

Hankiiiiiiiiiiiiieepankie, zonder dat je het weet, heb jij een belangrijke bijdrage geleverd aan de totstandkoming van dit proefschrift. Sinds jij op je vijfde al leerde lezen, en ik vele woorden later met tegenzin je voorbeeld besloot te volgen, heeft deze trend zich de rest van ons leven voortgezet. Ik heb jou, die altijd twee passen op me voorliep, altijd willen bijbenen (sowieso al een oneerlijke strijd omdat jouw benen letterlijk en figuurlijk twee keer zo lang zijn als de mijne). Maar op het gebied van promoveren ben ik jou in ieder geval voorgebleven! Of denk je dat je argument dat ik tien minuten ouder ben nu ook geldt!? Reuze bedankt in ieder geval ook voor de vele gezellige belletjes en je oprechte interesse in de voortgang van mijn promotie.

Ina en Erik, zus en broer, jullie ook bedankt voor jullie steun! Erik, jij bent altijd heerlijk nuchter en dat helpt ('Had je moar geen doctuhr moetn wordn'). Daarnaast ben je de meest vrijgevege persoon die ken. Heel erg bedankt nog voor de Wii die je samen met Hanke aan mij en Luuk hebt gegeven, waardoor we tijdens de laatste loodjes van het proefschrift de nodige ontspanning en lol hebben kunnen zoeken. Ina, dankzij jou is mijn interesse voor de studie psychologie en de stad Maastricht gewekt. Ik heb het altijd heel fijn gevonden dat mijn grote zus zo dichtbij woont. Dank je wel dat je mijn bezigheden op de universiteit altijd op de voet hebt gevolgd. Maar vooral heel erg bedankt voor de vele gezellige dingen die we samen, maar ook met Marit en Bente, ondernemen.

Oma, ook een paar woorden aan u. Ik vind het geweldig bijzonder hoe u nog met beide benen in het leven staat. Het feit dat u op uw 80<sup>e</sup> nog met computeren en internetten bent begonnen, is hier een duidelijk voorbeeld van. Ik geniet van onze gesprekken en het van tijd tot tijd discussiëren en steggelen ('Ik zal geen bloemen meenemen als u mij geen benzinegeld meer geeft'). Dank u wel voor uw betrokkenheid bij mijn werk en de rest van mijn leven.

Mijn 'schoonfamilie', Frans, Door en Stefan, jullie heel erg bedankt voor de eeuwige gastvrijheid en gezelligheid die bij jullie te vinden is. Jullie interesse in mijn werk, en nog zoveel meer, heb ik altijd erg gewaardeerd.

Luuk, jij hebt nog wel het meeste te verduren gehad en verdient de meeste lof voor je bijdrage aan mijn promotie. Jij hebt elke dag 'moeten' meeleven met mijn proefschriftperikelen. Zeker tijdens de laatste zware loodjes op een lege batterij, en dan ook nog in combinatie met een nieuwe baan, heb je heel wat opvliegendheid van me moeten verdragen. Gelukkig ken je mijn gebruiksaanwijzing ondertussen door en door en weet je precies goed te reageren. Je knuffels op de moeilijke momenten waren van onschatbare waarde. Ook al zei het je vaak weinig, toch heb je altijd proberen te begrijpen wat ik elke dag achter mijn beeldscherm heb uitgespookt. Je hebt me op de juiste momenten weten af te remmen, of juist te motiveren, vooral door te luisteren en alleen maar het beste voor me te willen. Je overspoelt me met onvoorwaardelijk veel liefde en zorg. Bij jou kan ik thuiskomen en volledig mezelf zijn en dat is het fijnste gevoel dat je me kan geven. Je maakt me erg gelukkig. Een oneindig dank je wel voor alles!



## **Curriculum Vitae**





### **Curriculum Vitae**

Maaïke Leeuw was born on 29 January 1979 in Hengelo, the Netherlands. She graduated from secondary school, Grundel Lyceum in Hengelo in the Netherlands, in 1997. She subsequently moved to Maastricht to study psychology at Maastricht University, the Netherlands. After finishing two additional voluntary clinical internships at the Academic Anxiety Centre and the Department of Neuropsychology of Vijverdal in Maastricht, she graduated in June 2002. She received her master's degree in cognitive psychology, with the main subject 'psychopathology'. In September 2002, she started as a PhD-student at the formerly known Department of Medical, Clinical, and Experimental Psychology (currently named Clinical Psychological Science) of Maastricht University. The studies included in this dissertation were carried out as part of her PhD research project. In addition to her PhD research, she worked voluntarily as a psychologist at the regional institute for mental welfare (RIAGG) Maastricht for one day per week from March 2004 until August 2005. Since August 2007, she is appointed as a psychologist in Rehabilitation Centre Hoensbroeck in Hoensbroek the Netherlands, where she participates in the multidisciplinary treatment of chronic pain patients.



## **Publications**



## Publications

- Leeuw, M.**, Goossens, M.E.J.B., de Vet, H.C.W., Vlaeyen, J.W.S. (accepted for publication, pending revision). Assessing protocol adherence and contamination in treatment outcome studies: A methodological illustration from behavioural medicine. *Journal of Clinical Epidemiology*.
- Leeuw, M.**, Goossens, M.E.J.B., van Breukelen, G.J.P., de Jong, J.R., Heuts, P.H.T.G., Smeets, R.J.E.M., Köke, A.J.A., Vlaeyen, J.W.S. (In press). Exposure in vivo versus operant graded activity in chronic low back pain patients: Results of a randomized controlled trial. *Pain*.
- Leeuw, M.**, Goossens, M.E.J.B., van Breukelen, G.J.P., Boersma, K., Vlaeyen, J.W.S. (2007). Measuring perceived harmfulness of physical activities in patients with chronic low back pain: the Photograph Series of Daily Activities - Short electronic Version. *The Journal of Pain*, 8(11), 840-849.
- Leeuw, M.**, Goossens, M.E.J.B., Linton, S.J., Crombez, G., Boersma, K., Vlaeyen, J.W.S. (2007). The fear-avoidance model of musculoskeletal pain: current state of scientific evidence. *Journal of Behavioral Medicine*, 30(1), 77-94.
- Leeuw, M.**, Houben, R.M.A., Vlaeyen, J.W.S. , Severeijns, R., Picavet, H.S.J., Schouten, E.G.W. (2007). Pain-related fear in low back pain: a prospective study in the general population. *European Journal of Pain*, 11(3), 256-266.
- Leeuw, M.**, Peters, M.L., Wiers, R.W. & Vlaeyen, J.W.S. (2007). Measuring fear of movement/(re)injury in chronic low back pain using implicit measures. *Cognitive Behaviour Therapy*, 36(1), 52-64.
- Leeuw, M.**, Vlaeyen, J.W.S. & Crombez, G. (2007). Fear of movement/(re)injury. In: R.F. Schmidt & W.D. Willis (eds.). *Encyclopedia of Pain*. Springer-Verlag, Berlin. ISBN: 978-3-540-43957-8 (Print) 978-3-540-29805-2 (Online).
- Leeuw, M.** & Vlaeyen, J.W.S. (2007). Fear of pain. In: R.F. Schmidt & W.D. Willis (2006). *Encyclopedia of Pain*. Springer-Verlag, Berlin. ISBN: 978-3-540-43957-8 (Print) 978-3-540-29805-2 (Online).
- Leeuw, M.**, Vlaeyen, J.W.S., Janssen, R., van Riet, S. (2007). Photograph Series of Daily Activities-Short electronic Version: free software internet download, user manual and installation manual.
- Leeuw, M.**, Vlaeyen, J.W.S., de Jong, J.R., Goossens, M.E.J.B. (2006). Exposure in vivo bij chronische lage rugpijn, werkboek en therapeutenboek. In: reeks cognitieve gedragstherapie. Boom Test Uitgevers. ISBN –10-9085063140 en ISBN –13-9879085063148

- Leeuw, M.**, Goossens, M.E.J.B., van Breukelen, G.J.P., Boersma, K., Vlaeyen, J.W.S. (2006). Het meten van de dreigwaarde van activiteiten in patiënten met chronische lage rugpijn: de Photograph Series of Daily Activities – Short electronic Version (PHODA-SeV). *Revalidata*, 134, 6-9.
- Houben, R.M.A., **Leeuw, M.**, Vlaeyen, J.W.S., Goubert, L., Picavet, H.S.J. (2005). Fear of movement/injury in the general population : Factor structure and psychometric properties of an adapted version of the Tampa scale for Kinesiophobia. *Journal of Behavioural Medicine*, 28(5), 415-424.
- Leeuw, M.**, Vlaeyen, J.W.S., Goossens, M.J.E.B. (2005). Exposure in vivo bij lage rugpijn. *De Psycholoog*, 40(4), 174-180.
- Leeuw, M.**, Vlaeyen, J.W.S., Goossens, M.J.E.B. (2005). Exposure in vivo bij lage rugpijn. *Beweegreden: vakblad voor oefentherapeuten caesar en mensendieck*, 1(3), 4 – 9 (reprint van artikel *De Psycholoog*).
- Geilen, M.J., **Leeuw, M.** & Hodiamont-Joosten, M.C.J.(2005). Graded activity volgens gedragsgeoriënteerde principes. In: P.U. Dijkstra, P. van Burken, J. Marinus, J. Nijs, C.P. van Wilgen. *Jaarboek Fysiotherapie Kinesithérapie*. Bohn Stafleu Van Loghum, Houten.
- Vlaeyen, J.W.S., de Jong, J.R., **Leeuw, M.** & Crombez, G. (2004). Fear reduction in chronic pain: graded exposure in vivo with behavioral experiments. In: G.J.G. Asmundson, J.W.S. Vlaeyen, G. Crombez (Eds.). *Understanding and treating fear of pain*. Oxford University Press, Oxford, UK.
- Leeuw, M.** & Vlaeyen, J.W.S. (2004). Een cognitief gedragsmatige exposure in vivo behandeling bij chronische rugpijn. *Revalidata*, 121, 26-31.